Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2023, Vol. 15, No. 5

The Lexico-grammatical Ecology of Media Arabic Genre: A Multi-dimensional Analysis

Zakariae Azennoud

Culture, Language, Education, Migration and Society's Research Laboratory

Mohammed V University, Morocco

E-mail: zakariae.azennoud@gmail.com

Received: October 1, 2023 Accepted: October 26, 2023 Published: October 27, 2023

doi:10.5296/ijl.v15i5.21376 URL: https://doi.org/10.5296/ijl.v15i5.21376

Abstract

Media Arabic is a linguistically independent genre of Modern Standard Arabic. There are a number of explanatory reasons accounting for a characterized, integrally structured media Arabic form. Much within our scientific grasp is the overall combination of lexico-grammatical traits, constituting a major part of the internal textual ecology. Within our reach, again, there is parsimonious literature hardly informing, or predicting, what situationally and linguistically a media Arabic text is. The vast body of research is intuition-governed and prescriptive. This paper is a concise exposition of a work endeavoring to fill out such a gap in research. It summarizes the findings of a large-scale project in Arabic corpus linguistics that empirically investigates media Arabic genre from a functional lexico-grammatical perspective, by using Multi-Dimensional framework. The model uncovers the parameters of variation in media Arabic and explores underlying co-occurring functions. In documenting MSA textual similarities and differences among ten written media subgenres: Politics, Economics, Society, Sports, Science, Culture & Arts, Technology, Women, Health & Medicine, and Religion, Exploratory Factor Analysis was run and resolved within a six-factorial structure, the raw output to be subsequently interpreted as six-dimensional basis for the variable use of Arabic across the ten subgenres.

Keywords: Media Arabic genre, Lexico-grammatical description, Multi-dimensional framework, Dimensions of variation

1. Background

Genre is viewed, in literary and scientific disciplines where it has been used, so differently that it raised a critical conceptual discussion. This issue of terminology expanded far more seriously when then-overlapping notions, chiefly among which are *text type*, *style*, and

register partook of the general interests of linguists (e.g. Paltridge, 1996; Swales, 1990). The characterization of 'media Arabic' as a distinct genre thus has been inevitably born out of a synthetic, coherent review of contextualized literature. The review allowed revisiting the researcher's expertise in the field of Teaching Arabic as a Foreign Language (TAFL) and proceeding with this experimentally-oriented subject matter in a systematic fashion of research.

Before turning to the pivot of the present study, a concise sketch, and succinct as it might be required, of media-Arabic-as-a-genre conception is mandatory. From a functional perspective, genre reflects the cultural context within which texts are constructed (Martin, 2001; Swales, 1990). Imported to linguistics from literary and rhetorical analysis (Nunan, 2008), its borders were redrawn with respect to universal classification criteria. In modern corpus linguistics, genres are considered as the categories of texts chosen on the basis of fairly easily definable external parameters (Biber, 1988; Lee 2001). Broadly, there is an established consensus among scholars (Biber, 1988; Biber & Conrad, 2009; Lee, 2001; Paltridge, 1996; Sinclair & Ball, 1996; Steen, 1999) that genre is based on external, non-linguistic, traditionally evident criteria, conventionally and culturally recognized group of texts, characterized according to the intended audience, purpose, topic, and activity type.

Media Arabic (MA) is, therefore, conducive to such unanimous concept of genre, for the external elements of it are easily recognizable, be it spoken such as radio, or written such as newspapers. MA also reflects style and register, yet on a limited scale. Michael Halliday and Ruqaiya Hasan (1985) define *register* as a particular configuration of field, tenor, and mode. It is associated with particular communicative purposes and situational contexts of texts; whereas *style* mirrors the internal properties of individual texts or the language used by individual authors (Biber & Conrad, 2009; Lee, 2001). Here neither register represents MA, nor does the concept of style. To avoid any depriving interrelated matter of inquiry, Biber and Conrad (2009) as well as David Nunan (2001) emphasize that genre and register are complementary rather than alternative ways of textual analysis. Genre, in addition to the cultural and conventional aspects, allows the research to move from documenting the external landscape of the form of language under question to exploring an internal ecology, a multi-dimensional network of functional relationships in regard to sub-categories of texts (subgenres), determined by underlying lexico-grammatical factors. As far as this study is concerned, it is this latter linguistic exploration to which the present research endeavor is confined.

Besides the typological indicators of genre stated above, it is assumed that MA, growing as a contemporary genre of Modern Standard Arabic (MSA), has been brought about by various factors. On a regular basis, journalists officially consult Arabic language academies about language choices and ask for recommendations. The academies examine suggestions and publish the output which draws on comparative research and leaves room for innovation. The continuous reference to these institutional bodies by the press to make decisions about language use unifies the genre. Additionally, journalism, as an art, is a functional practice, a tool not aesthetic in its own right, but tends to reach a large proportion of readership (Al-Bakkae, 2010). MSA thus satisfies such a widely interactive need to a large degree. By

the same token, the criteria of content editing in media regarding style, word selection and usage, clarity, and simplicity determine what Mahmoud Adham calls the 'press level' (n.d). The artistic press level is distinguishable from the scientific level which requires a technical jargon, or the literary level that is dependent on a critiquing taste and emotional involvement (Adham n.d.). Furthermore, it is inescapable a law of development in natural languages the effects of use and disuse, "that the use strengthens certain parts [or structures], and disuse diminishes them and that such modifications are inherited" (brackets added), to quote Charles Darwin (1859: 133). Arabic is certainly no exception (Chejne, 1969; Ostler, 2006; Saleh, 2007; Zaydan, 1988). In this cited research, Abdul Rahman El-Haj Saleh, for example, emphasizes the point in comparing between oral and written forms of MSA. He (2007) believes that the extensive use of language is more exposed to structural change. Newspapers exemplify this standpoint as they produce language content on a daily basis and, with the advent of online papers, the content might be updated several times a day. The review also remarkably reveals that the Arabic language of the media is variable in regard to subgenres (e.g. politics, society, arts, etc.), the sub-systemic categories of texts externally situationally and linguistically distinguishable.

Based on an updated review, there has been a shortage of scientific research investigating media Arabic on a deep linguistic level; concerning its subgenres, far more pressing insufficiency is witnessed. Except the PhD thesis of Zainur Rijal Abdul Razak (2011), the dominant body of research centers on the language of Arabic media as a part of MSA; although researchers (Al-Bakkae, 2010; Belaid, 2006; Hammou, 2010; Sharaf, 1991; among others) admit that MA is a complex and distinct system of integrated linguistic structures, they account for such complexity either by shedding light on deviations from standard rules in the media use of language, where a prescriptive orientation is highly persistent, or by shouldering an intuitive approach more of an explanatory burden than it could bear. The status of the existing relevant research, hence, does not cater for the multi-fold need of Arabic applied linguistics. In a nutshell, our incomplete understanding of the linguistic characteristics of MA genre, the lack of adequate background research to enable solid theoretical predictions about the similarities and differences among its subgenres, and a dominant intuitive tendency about functional contexts in which co-occurring sets of structures work, in an era where computer-readable data analysis is fairly possible more than any time ago, are the main reasons the present study came out into existence.

It is a corpus-based factorial study that tends to explore the underlying dimensions of variation in MA genre. This study answers four research questions:

1) What are the sets of lexico-grammatical patterns that co-occur in MA genre?
2) How are these sets distributed across MA subgenres?
3) To what extent do MA subgenres differ from each other within a multi-factorial structure?
4) What are the appropriate functional interpretations that can be advanced to account for the dominant factors of variation in MA genre?

The answers to these questions seek to fulfill two main objectives:

1) To document the lexico-grammatical characteristics of MA genre and subgenres.

2) To establish a systematic multi-dimensional reference of the lexico-grammatical similarities and differences among the subgenres of MA that has potential implications for many applied areas, among which is syllabus design.

To investigate the lexico-grammatical ecology of media Arabic genre, Multi-dimensional framework is adopted. The method, including the statistical technique Exploratory Factor Analysis (EFA), is laid out in the section below.

2. Multi-dimensional Method

Multi-Dimensional (called also *Multi-feature*) model was developed by American applied linguist Douglas Biber in the mid-1980s. Since that time (Biber, 1986a, 1986b, 1987; Biber, & Finegan, 1988), Multi-dimensional (MD) framework has increasingly shown unquestionable success in analyzing the English language, its varieties, style, register, and genre. Much remarkably, the model was not limited to English, its application extended to other languages, including Nukulaelae, Korean, Brazilian Portuguese, and Somali. Arabic language, however, as demonstrated by navigation in widely searched publication databases, has not taken advantage of it but in a very limited way. Very limited, because there is only a single study (Mohamed & Hardie, 2019) that attempted to construct a typology for the Arabic language, by using cluster analysis, similar to Biber's *A Typology of English Texts* (1989). The option of MD framework was thus catalyzed by two incentives. First, it came as a strong impetus, in the absence of a rich MD body of research on Arabic. Second, the MD methodology is highly conducive to answering the research questions raised above in a feasible, productive way.

As interdisciplinary as it will be practically manifesting across this paper, the full MD method functions within two applied areas, corpus design and computational linguistics, merged by Leech (1992) in a coinage termed *computer corpus linguistics*. The overall MD methodology as is adopted in the present study can be blueprinted in a six-step research design:

1) Compilation of a representative media Arabic corpus.

2) Building of the coding lexico-grammatical scheme to be included in the analysis.

3) Lexico-grammatical analysis of the corpus, using appropriate computer programs (mainly Farasa software program), supported by manual revisions, to obtain frequency counts of each feature included in the coding scheme.

4) Analysis of the co-occurrence patterns among lexico-grammatical features, using EFA of the frequency counts.

5) Computation of factor scores.

6) Functional interpretation of factors as underlying dimensions of variation in MA with reference to the literature written on the features under discussion.

Exploratory approach, using EFA, was preferable to confirmatory one for one reason, emanating from the review conducted about studies on MA. It is common among researchers that the confirmatory approach is implemented as a complementary alternative to investigate

the scientific generalizations about the existence or non-existence of particular linguistic attributes. When a genre under analysis lacks basic scientific judgments, quantitatively replicable and qualitatively valid, about its linguistic structures and functions, an exploratory approach is indispensable. Now a comprehensive critical review informs us that MA lacks a solid theoretical basis, a scientific corpus-based linguistic characterization of the MA genre, which justifies an exploratory study interest.

2.1 Media Arabic Corpus

Research data is drawn from the *media Arabic corpus*. The dataset that include specific linguistic information about the corpus functions as the basis for subsequent statistical technique chosen by the researcher as a method of analysis (EFA). The construction of media Arabic corpus follows sample corpus approach (McEnery & Hardie, 2012). The corpus includes print and electronic press materials (Figure 1, right). Articles were selected from twelve newspapers and news websites whose headquarters are situated in ten countries.

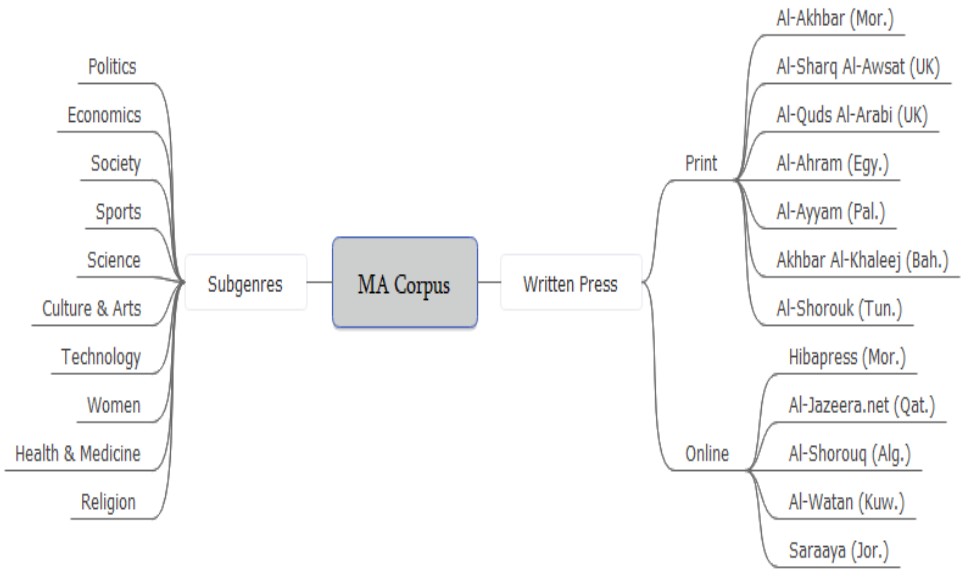

Figure 1. Sources of media Arabic corpus

Classification of texts and labeling of subgenres, in terms of corpus representativeness (Atkins et al., 1992; Biber, 1993; Brezina, 2018; Leech, 2007; Sinclair & Ball, 1996), were carefully studied. The classification was decided on ten subgenres: *Politics, Economics, Society, Sports, Science, Culture & Arts, Technology, Women, Health & Medicine, and Religion* (Figure 1, left). Texts were collected in a five-month period, extending from January 1st to June 18th, 2020. Overall, the corpus contains approximately 106 070 words, of 212 press articles (Table 1).

Table 1. Distribution of texts across ten subgenres

Subgenres	No. of texts
Politics	21
Economics	23
Society	24
Sports	21
Science	16
Culture & Arts	25
Technology	19
Women	22
Health & Medicine	23
Religion	18
Total number of texts	212
Approximate no. of words	106 070

The number of words in each text ranges approximately from 500 to 1 000 words (to put it into central tendency, the word count mean reaches 488.14). This is the ideal number of words that is recommended (Biber, 1988; Biber & Conrad, 2009; Friginal & Hardy, 2014). In relation to lexico-grammatical distribution of individual occurrences, the texts that violate the principled number of words were balanced using a common calculation called *normalized/relative* frequency, which is obtained by dividing the number of instances (occurrences) in a particular text by the total number of words of that text, then multiplying the sum by a basis of 1 000.

2.2 Coding Lexico-grammatical Scheme

Lexico-grammatical coding scheme is a basic analytical input in MD analysis. To annotate the MA corpus, it was crucial to compose a list of comprehensive lexico-grammatical features. Initial scheme was drawn from a careful reading of numerous texts from the corpus, survey of books and textbooks in the MA, and review of corpus-based studies. The list had encompassed more than 70 features. After testing the scheme, a number of features did not hold salient; some of them were even extremely poor in the whole corpus. It is worth noting that weak linguistic salience makes a schematic choice liable to invalidity. To maintain high probability of valid, reliable, and practical application, the scheme underwent modifications that involved the removal of some features and the inclusion of others. The rebuilding of the linguistic scheme was also founded on piloting it on other corpora and initial EFA results. A finalized scheme used to annotate and analyze the MA corpus includes 50 lexico-grammatical features, as shown in the following:

A. Tenses

1. past tense
2. present tense
3. future with *sa* particle
4. future with *sawfa* particle
5. future with *kay* particle
6. future with *lan* particle
7. future with *'alā* or *'amā* particles
8. future with *'in* or *law* particle

B. Verbs

9. imperatives
10. passive form
11. verbs of imminence (*al-muqāraba*)
12. verbs of probability (*ar-rujḥān*)
13. *kāna* and sisters: *kāna* and *mā zāla*
14. *kāna* and sisters: *mā dāma*
15. *kāna* and sisters: *'aṣbaḥa*
16. *kāna* and sisters: *bāta* and *ḍhalla*
17. *kāna* and sisters: *ṣāra*

C. Pronouns

18. attached pronouns
19. detached pronouns
20. first-person pronouns
21. second-person pronouns
22. third-person pronouns

D. Particles

23. negation
24. *'inna* and sisters: *'inna*
25. *'inna* and sisters: *'anna*
26. *'inna* and sisters: *ka'anna*
27. *'inna* and sisters: *lākinna*
28. *min* preposition
29. *'ilā* prepostion
30. *'an* prepostion
31. *'alā* prepostion
32. *fī* prepostion
33. *al-bā'e* prepostion
34. *al-kāf* prepostion
35. *al-lām* prepostion
36. *fa-* of continuation
37. *hasaba*
38. *wifqan li*

E. Adjectives

39. adjectives

F. Adverbials

40. adverbs of place
41. adverbs of time

G. Nouns

42. proper nouns
43. definite nouns
44. numbers

H. Other grammatical styles

45. interrogative sentences
46. near demonstratives
47. remote demonstratives
48. relative sentences
49. exceptional sentences
50. conditional sentence

The annotation of the MA corpus was carried out using Farasa software program (Abdelali et al., 2016), supported by manual revision as well as the contextual reading of features' occurrences.

3. Results

A *factor* explores the strength of textual lexico-grammatical relations among subgenres. It reflects correlational structure of lexico-grammatical occurrence. The lexico-grammatical features that tend to co-occur together in a consistent way, as will be shown below, are likely to be structured in a specific factor of variability, encompassed in a particular dimensional text type, as a consequence. By and large, the results of the study are provided by initial and rotated EFA.

3.1 Exploratory Factor Analysis: Initial and Rotated

Initial EFA provides a scree plot and an adequacy extraction measure. The scree plot, shown in Figure 2, is run before a rotation method, on the purpose of determining the adequacy of extraction and to identify the likely number of factors in the solution.

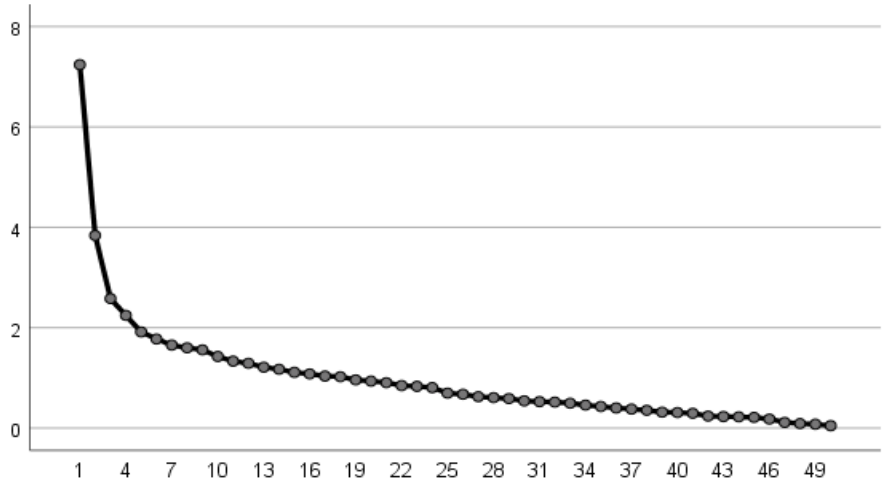

Figure 2. Scree plot of textual factors

The diagram plots a 50-factor rising line vis-à-vis almost 7.3 eigenvalues. Increasingly, each factor captures a certain amount of variance. To the naked eye, the movement of the line helps focus on particular curves; a point below eigenvalue 2 is plainly visible. The step of factor extraction, however, must be careful. Despite overall extraction of factors reaches 0.703 using Kaiser-Meyer-Olkin Measure, which is considered adequate (Horn & Martin, 2008). Partial factorial extraction might produce weak factor structure, ensuring limited variables and poor loadings.

Back to Figure 2, plotted factors (horizontal) scaled in eigenvalues (vertical) suggest three possible solutions. Solution one is the extraction of all factors which have a total eigenvalue above 1 (e.i. 18 factors). Solution two is to extract factors with regard to the movement of the line in the plot, that is, where the curve flattens out below 2. Solution three is to extract factors that explain great variance among all given lexico-grammatical features; in this case, it is difficult to judge, because, except the first five, all factors descend in a very approximate range value. To avoid inexact judgment about factorial extraction solution, it is thus compulsory to *rotate* the EFA test and uncover the results in a specific, readable output.

Rotation method used in the present study is Promax. It negotiates all solutions above and chooses the extraction that produces the best factor structure. Good structure is measured with factor loadings above 0.30, fewer cross loadings, and a larger number of variables per factor (Biber, 1988; Costello & Osborne, 2005; Gorsuch, 1983). Following the Promax rotation method, solutions one and two harmonize in one choice, extracting 6 factors that account for about 40% of the total variance, taken into granted the plot view, along with confirmatory factorial structure that include more functionally interpretable correlations. The following table highlights the six extracted factors, accompanied by their eigenvalues and variance amounts.

Table 2. Six factors extracted

Factors	Initial Eigenvalues			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total
1	7.238	14.477	14.477	7.062
2	3.837	7.673	22.150	4.065
3	2.579	5.159	27.309	3.132
4	2.245	4.490	31.798	2.840
5	1.916	3.832	35.630	2.575
6	1.774	3.548	39.179	1.999

Rotated six-factorial structure, which will be presented in the next section in turn, based on the extraction solution above, is derived from a correlation matrix of all variables. Yet a typical brief presentation of co-occurrence statistics, unified and compact, can only be done through a well-integrated structure, constructed by Promax-rotated EFA.

3.2 Six-factorial Structure and Lexico-grammatical Distribution

In the final factor solution, rotated in Promax, the factorial structure results in two-opposite line loadings, with positive and negative weights. The loading of a feature represents to what extent the frequency variation of that feature correlates with the overall variation of a factor. In other words, it reflects the co-occurrence relations between the feature under question and the factor as a whole. Generally, salient loadings are included as statistically significant, coming in handy in the functional interpretation of the factor results. As noted earlier, scholars unanimously see that loadings having a value above .30 must be interpreted as significant part of a factor. In this way, such norms are followed as a major procedure of selection, but few features below .30 are also included for two reasons. First, their absolute values do not exceed .30 (8 features). Second, their values, which are very close to the recommended loading, are deemed helpful in understanding more thoroughly the lexico-grammatical co-occurrence of the dimension in the phase of discussion. Table 3 presents a summary of the six factors derived from rotated EFA of 50 lexico-grammatical features in MA texts.

Table 3. Summary of the six factors derived from EFA

FACTOR 1		FACTOR 2	
Positive features:		Positive features:	
interrogative sentences	*.90*	*attached pronouns*	*.81*
remote demonstratives	*.81*	*third-person pronouns*	*.76*
min preposition	*.78*	*(present tense)*	*.55*
detached pronouns	*.77*	*future with lan particle*	*.49*
'an preposition	*.73*	*verbs of imminence*	*.46*
al-lām preposition	*.70*	*verbs of probability*	*.45*
ka'anna particle	*.62*	*(detached pronouns)*	*.45*
first-person pronouns	*.59*	*(kāna and mā zāla)*	*.40*
albā'e preposition	*.55*	*'inna particle*	*.36*
relative sentences	*.54*	*(fa- of continuation)*	*.36*
kāna and mā zāla	*.51*	*mā dāma*	*.18*
'anna particle	*.41*		
(near demonstratives)	*.36*	No negative features	
(attached pronouns)	*.35*		
'alā preposition	*.35*		
(past tense)	*.34*		
wifqan li	*.34*		
conditional sentences	*.31*		
future with sa particle	*.26*		
'ilā preposition	*.21*		
No negative features			

FACTOR 3		FACTOR 4	
Positive features:		Positive features:	
second-person pronouns	*.89*	*present tense*	*.59*
imperatives	*.84*	*(negation)*	*.47*
fa- of continuation	*.71*	*near demonstratives*	*.37*
future with 'in particle	*.51*	*(fa- of continuation)*	*.35*
(present tense)	*.30*	*('inna particle)*	*.30*
		al-kāf preposition	*.19*
Negative features:			
(adjectives)	*-.37*	Negative features:	
		numbers	*-.59*
		past tense	*-.49*
		proper nouns	*-.49*
		adverbs of time	*-.44*
		fī preposition	*-.34*

FACTOR 5		FACTOR 6	
Positive features:		Positive features:	
definite nouns	*.67*	*'aṣbaḥa*	*.55*
adjectives	*.62*	*(first-person pronouns)*	*.49*
passive form	*.35*	*future sawfa particle*	*.35*
		(kāna and mā zāla)	*.33*
Negative features:		*lākinna particle*	*.32*
negation	*-.5*	*(verbs of probability)*	*.31*
(first-person pronouns)	*5*		
(lākinna particle)	*-.3*	Negative features:	
(conditional sentences)	*5*	*ṣara*	*-.39*
future with kay particle	*-.2*	*adverbs of place*	*-.26*
	9	*exceptional sentences*	*-.25*
	-.2	*(third-person pronouns)*	*-.23*
	9	*bāta or ḍhalla*	*-.20*
	-.2	*hasaba*	*-.20*
	4	*future with 'alā or 'amā particle*	*-.13*

Positive and *negative* weights are opposite factorial categorizations of correlational structures. Positive features in a particular factor normally co-occur in certain subgenres of MA whose texts, in turn, do not include features that are found negative in the same factor. For instance, there is poor (negative) co-occurrence of negation, first-person pronouns, *lākinna* particle, conditional sentences, future with *kay* particle in texts in which definite nouns, adjectives, passive form significantly co-occur (positive in factor 5). Likewise, texts in which the negative features frequently co-occur have poor co-occurrence of positive features. Another notice to make is that the lexico-grammatical features between parentheses are used twice. They are considered as secondary variables because their loadings are acceptable (above 0.30 or below -0.30) help in the interpretation of factors as dimensions of variation in the MA genre. The co-occurrence of lexico-grammatical features within the positive and negative poles, as found by EFA, implies a two-opposite-extreme dimensional spectrum. In fact, placing subgenres of MA on dimensional scales goes hand in hand with the functional interpretation of factors as underlying dimensions of variation in the genre, a part of the study to be presented in detail in the following section.

4. Discussion

There are six factors drawn from EFA of MA genre. In the discussion that follows, the variation in MA genre is, therefore, *six-dimensional* from a MD perspective. Each factor is discussed independently, with reference to rigorous synthesis of literature written on the lexico-grammatical features under discussion.

4.1 Interpretation of Factor 1

Factor 1 includes 20 lexico-grammatical features (see Table 3). All of them have positive loading weights. Based on the review of literature analyzing features under question, a number of common functions are synthetically explored. The use of interrogatives convey a demand

for certainty and visualization, as stated by scholars like Sibawayhi (1988 [1975]) and Ar-Rajihi (1998 [1972]). Interrogatives generally prompt a response from either a physical or virtual recipient. In this context, Marckworth and Baker (1974) as well as Biber (1986a) state that questions indicate a concern with interpersonal functions. The involvement with the addressee is framed not only by the strong co-occurring patterns of interrogative and demonstrative nouns, but also, implicitly, by the highly-loading four prepositions: *min*, *ʿan*, *al-lām*, and *al-bāʾe*. Indeed, these prepositions express concrete indication and presupposition (Al-Jurjani, 1982; Al-Muradi, 1992; Sibawayhi, 1988 [1975]), used to focus the reader's attention on location or time in the speech situation (Igaab & Tarrad, 2019). As a marker of discoursal coherence (Halliday & Hasan, 1976), presuppositional function is employed by the author to set situational boundaries for a particular event setting. The two other prepositions, *ʿala* and *ʾila*, included in this factor in a modest correlation with other prepositional structures, share a location-specific function (Al-Shamsan, 1987; Ibn Faris, 1971; Sibawayhi, 1988 [1975]), a communicative property found, albeit in varying degrees, in all the above-stated features. Another function to synthesize is the synta-semantic dependence, emerging in multifold manifestations.

Remote demonstratives, used as noun substitutes (Ryding, 2005), depend on context in their interpretation, the same dependent-feature aspect deemed inherent in detached pronouns, (Igaab & Tarrad, 2019). One of the purposes for which remote demonstratives and detached pronouns – along two other features appearing in this factor, *kaʾanna* and *ʾanna* particles – are employed is what Eid (1983), Belnap (1986), and Ryding (2005) consensually refer to as emphatic meaning of language. *kaʾanna* particle, detached pronouns, merged with relatives pronouns in the same texts, however, indicate broad, relevant-to-other-features functions, namely specification, restricted addition, extension (Beaman, 1984; Winter, 1982), integration (Salim & Ahmad, 2015) and different sorts of references (Mulyadi, 2021; Ochs, 1979). Again, the majority of them reflect dependence of syntactic units on each other, within specific situations, in order to generate a coherent, comprehensible discourse.

The use of pronouns is essential in any texts. The attached pronouns that seem to co-occur reasonably in this factor incorporate the first person. The presence of the first speaker, to link it appropriately to the use of other features, conveys immersive experience, in that the teller of the story knows closely the events he relates and, more concretely, the ideas he captures. As a marker of proximity, as described by Al-Tamimi (2016), the first-person voice draws the readers into the subgeneric narrative, making them feel they are part of the story, somehow involved in the subject matter that is transmitted.

Factor 1 exhibits a notable absence of tenses. While past tense is considered a secondary explanatory variable, for being parenthesized and having a weight .34 and future with *sa* particle having insignificant weight in comparison with others, it is only *kāna*-and-*mā zāla* feature that shows high significance. Lack of tenses implies shortage in verbs. In this factor, instead of verbs, prepositions and fixed particles are pervasive. Most of the later lexico-grammatical features share the functions of time and place-relevant content. Time and place are principal properties of situation. Situational description usually includes information about when a particular event occurred and where it took place. *kāna*-and-*mā*

zāla feature itself can be regarded as a marker of situational expression as it carries the continuous aspect of action.

Features co-occurring in Factor 1 initiate, so to conclude, macro-cohesive syntagmatic dimension. That is, they link the units of meaning on a relatively large syntactic scale through particles that bear meanings not in their own lexicon but in their reference to near or remote words (Igaab & Tarrad, 2019; Mulyadi, 2021; Quirk & Greenbaum, 1973), occurring in a situationally-governed context. In such an environment, lexico-grammatical patterns such as relative, demonstrative, and conditional sentences are very active. Based on this discussion, Factor 1 is thereby interpreted as dimension of *Situation-dependent reference*. The subgenres of MA are scaled in this dimension as shown in Figure 3.

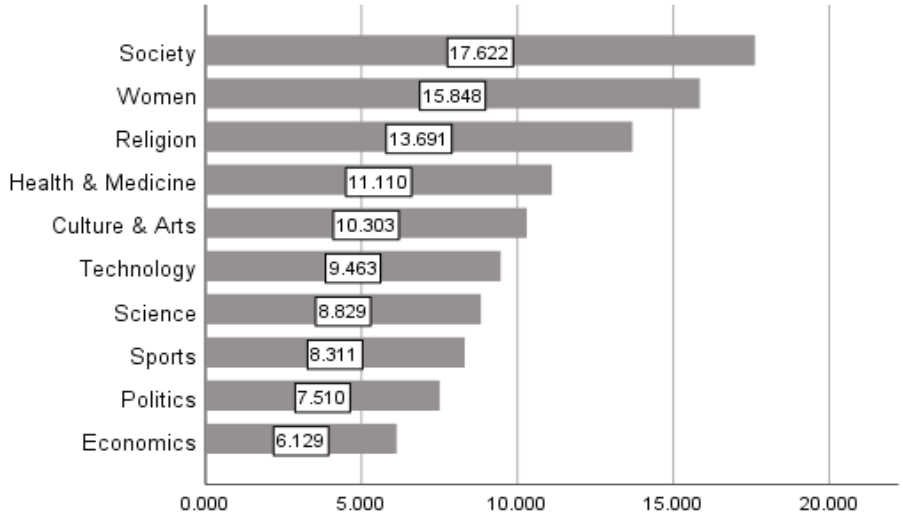

Figure 3. Dimension of *Situation-dependent reference*

All the ten subgenres of MA pertain to the positive part of *Situation-dependent reference* dimension. As an overall statement, it can be concluded that MA genre is firmly linked to the situational discourse. The extent of connection of each subgenre can be deduced from the place the subgenre has on the scale. At first glance, the graph above does not display a big disparity among the subgenres in respect to the *Situation-dependent reference* dimension. However, there are few of them that are posited in highly positive points, such *Society*, *Women*, and *Religion* subgenres, whereas *Economics*, *Sports*, and *Science* are not as very strongly relevant to the dimension as the former ones, implying a number of lexico-grammatical differences.

4.2 Interpretation of Factor 2

Factor 2 includes 11 positive features (Table 3). Similar to Factor 1, there are no features with negative loadings. Several common functions are concluded from the studies and analyses on the features co-occurring in Factor 2. Harmonious functional relationships are explored

among the co-use of present tense, future with *lan* particle, verbs of imminence, verbs of probability, and *ʾinna* particle, operating within a clearly defined frame of certainty, imminence, and reciprocal arrangement of a series of actions (Al-Mubarrad, 1967; Al-Sirafi, 2008; Al-Zobaidi, 1971; Ibn Faris, 1971; Sibawayhi, 1988 [1975]). In line with the classical grammar, Ryding (2005) and Hassan (2010) confirm that *ʾinna,* a marker of truth-intensifying expression, is used in MSA to introduce or report official speech. Among its chief uses, present-tense verbs deal with topics and actions of immediate relevance (Biber, 1988; Grabe & Biber, 1987), and thus also express facts, habits, and continuity in an occurring course of action (Al-Mallakh, 2009; Ibn Hicham, 1985; Hosni, 1997; Ostrabadhi, n.d.). In parallel, future with *lan* particle suggests, as described in Dendane & Dendane's analysis (2012), absolute negation with no possibility of future realization. It conveys assertive information about the non-occurrence of one or a series of interconnected actions. Attached and detached pronouns combined are employed in the third-person mode. Pronouns in general, as Al-Suyoti (2006) and Anis (1994) state, are used a way of economy. Third-person pronouns tend to hide the identity of the subject, avoid repetition, and create broad links between clauses. Investigating them as nodes in the corpus, they are also found to voice assertive style of expression.

While the aspect of verbs in the past is perfect, it is imperfect in present and future tenses (Al-Dobaian, 2018). The latter indicates unfinished action, incomplete events as the time involved has not elapsed yet. Therefore, in addition to assertive discourse, the features that co-occur in Factor 2 express open dynamic content. Openness of actions is internally cemented by the use of *fa-* of continuation. To Hussein (2011), *fa-* of continuation is an internal marker of connection and cohesion. It links words, phrases, and clauses in a meaningful way. Resided at the beginning, it is frequently used with verbal sentences, where it alludes to an extended point, a turn in conversation, or a forthcoming conclusion.

Open dynamic content is supported by the overall verbal landscape of the co-occurrence patterns. Unlike Factor 1, this factor translates a verbal-driven content. Major features consist of verbs (of imminence and probability), tense (present), and verb forms (*kāna* and *mā zāla* and *mā dāma*). Further, the particle *lan*, unlike *lā*, can negate only verbs, rendering it a verbal indicator. Texts that favor *lan* over other negation particles are presumed to be involved in a dynamic setting, influenced by a lively context. Similarly, *ʾinna* operates like verbs, conveying verb-like meanings in the word order they operate on. Both features reinforce the positive or negative continuity of the action, which is also crystallized by the use of *kāna* and *mā zāla* and their sister *mā dāma (*Al-Matlabi, 1986; Ibn Aqil, 1997; Ibn Yaish, 2001). Based on this assessment, Factor 2 is labeled dimension of *Open dynamic expression of assertion* (Figure 4).

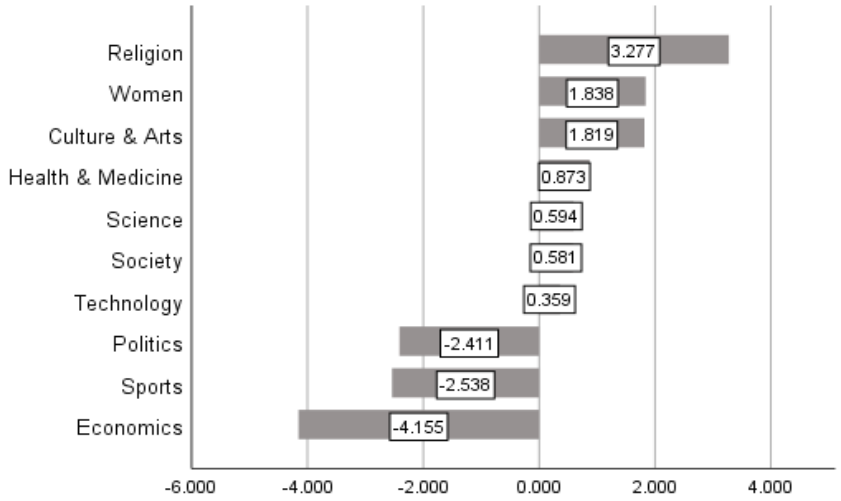

Figure 4. Dimension of *Open dynamic expression of assertion*

Religion subgenre represents the best example of dimension 2, followed by *Women* and *Culture & Arts*. *Health & Medicine*, *Science*, *Society*, and *Technology* are also positively located on the scale, without gradual degree of disparity. *Economics*, *Sports*, and *Politics* are negatively connected to this dimension.

4.3 Interpretation of Factor 3

Factor 3 includes 5 positive features and one negative feature (Table 3). The positive features share several functions. Maintained by the presence of both the addresser and addressee, second-person pronouns and imperatives indicate interactive relationships (Igaab & Tarrad, 2019). Because the role of the addressee, though hypothetically, is necessarily involved (Biber, 1986a; Marckworth & Baker, 1974), they are frequently used in interactive discourse. Imperatives express a request or prohibition that is to occur after the time of speech. Addressing the second-person pronouns in all situations, the imperative reflects interactivity, immediacy, and temporal externalization. Temporal externalization refers to expressions within which time is undefined or external. In a similar vein, present tense, as Biber (1988) observes, removes focus from any temporal sequencing. Highly loading positive features, cohered through ties like *fa-* of continuation, also encourage social proximity, shared destination, and respect (Al-Shahri, 2004).

In spite of the strong verbal attendance in this factor, the functional structure is either flexible or indefinite. In interactive discourse in general, the concept of time typically does not hold great importance. Progressive, occurring interpersonal relationships are rather sustained instead. The frequent co-occurrence also reinforces this deduction, in that it furnishes a time-being environment in which the realization of action, *jawāb* (response) to particular demands, is not specified. Still, actual meanings in the use of three features, imperatives (Basharat, 2017), a future with *ʾin* particle (Al-Rajihi, 1998 [1972]; Ibn Malek, 1967), and to a lesser degree simple present (Sibawayhi, 1988 [1975]), are realized in the future.

There is a kind of inter-dependent contract between the writer and the reader, by which the message of the text is revealed and the destination of the action work within common knowledge context. Structured in a future wish of two or more clauses, the features of Factor 3, especially future with *'in* particle, are most likely found in inter-connected discourse, in which fragments of the written text are interrelated. Interconnected discourse corresponds to the informative style, in that pieces that constitute the text argument are coherently composed, and new information are delivered in a fleetingly plain way.

To recapitulate, the four features, second-person pronouns, imperatives, future with *'in* particle, and present tense, do express verbal actions, but it is not clear the time, or even the location, in which the action exactly occurs. On the other side, immediacy and interdependent conveyance-comprehensibility of the text are retained, beside interactivity and immediacy. The negative weight represented by the use of adjectives complements this interpretation. Adjectives, by contrast to the the positive group, are about elaborated, detailed presentation and expansion of information (Biber, 1988; Mohamed & Hardie, 2019; Naknawa, 2020). However, there is no linguistic co-occurrence in the negative pole; the occurrence of only a signle feature is not sufficient for giving it a function as an oppositional extreme in the dimensional scale. This factor is, therefore, labeled dimension of *Interactive writer-reader interdependency discourse*. Subgenres are scaled on this dimension as presented in Figure 5.

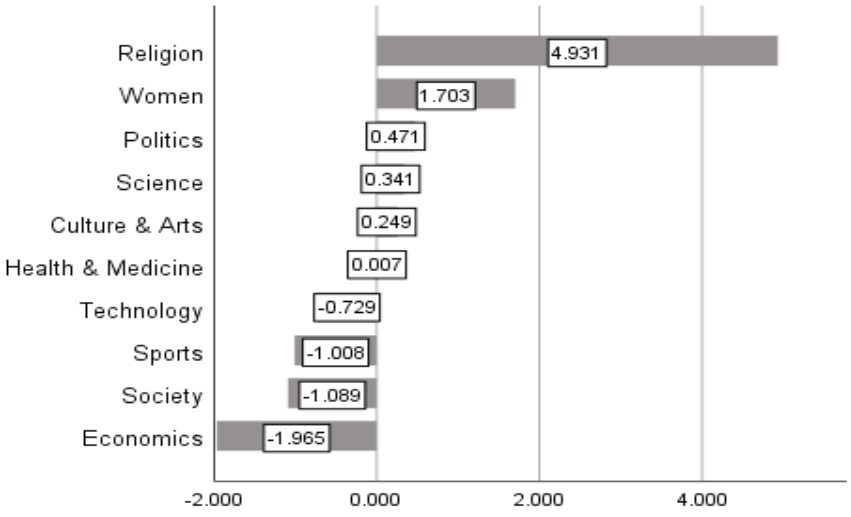

Figure 5. Dimension of *Interactive writer-reader interdependency discourse*

Some subgenres that rank high on dimension scales 1 and 2 (Figure 3; Figure 4) appear to have minimal, medium, or negative significance on this dimension. *Society*, *Technology*, and *Sports* are three cases in point. However, *Religion* still maintains a consistent position, with *Women* also holding a relatively prominent one.

4.4 Interpretation of Factor 4

Factor 4 includes 6 positive features and 5 negative features (Table 3). Building on the studies

of the functional use of the features under discussion, there exists a combination of shared communicative functions. Negation is a fundamental mental pattern, universally employed to express the non-existence or denial of action, object, or feeling. Languages make use of various strategies of it to convey negative meaning (Dendane & Dendane, 2012). Negative particles are used to negate actions in the past, present, or future. Some of them, as stated by Sibawayhi (1988 [1975]), are used to negate the present and future simultaneously. The common one is *lā*, a present-associated particle. The present tense is highly frequent in this factor. According to grammarians (Al-Mallakh, 2009; Ibn Hicham, 1985; Ostrabadhi, n.d.), the present tense, as mentioned earlier, points outs to habitual, factual, and continuous actions. Abunowara (2005) notices that specifically among other forms of language use, mass media frequently use simple present to describe events that are almost certain to happen. Back to negation, scholars like Al-Zamakhshari (1974) argue that *lan* contains certainty of negation in the future. In the same context, Al-Horais (2017) notes that *laysa*, a particle that can be preposited before verbs, nouns, and adjectives, is also found to be frequently used in the majority of structures as a focusing adverb. Texts using negative particles affirm the action regardless of its negative state, implying an abstract factual style.

The occurrence of events within a factual, continuous, and immediate context is functionally enhanced by the roles played by near demonstratives, as they refer to close objects, and *ʾinna* particle being a truth-emphasizing element. Holes (2004) attributes adjectival, presentative, and emphatic functions to the use of demonstratives. *al-kāf*, by contrast, serves as the major function of comparison (Abbas, 2009; Ibn Aqil, 1997; Mouawad, 2008; Zoubae, 2016). Co-occuring with the above-stated features, *al-kāf* indicates the abstract aspect of factuality. It allows imaginative experience by which the reader creates a mental picture to compare an object's descriptions, such as quality, degree, and size. The overall interpreted image summing up the discussion of the positive features exhibits factuality, certainty, and abstractness. The interpretation remains relative until it receives a complementation from the negative group.

The negative features express opposite functional style, yet only in terms of abstractness. The negative features are deemed concrete indicators of factual expression of language. From the definitions advocated by Sibawayhi (1988 [1975]), Ibn Mandhour (1968), and Al-Farahidi (1998), numbers expose statistical knowledge. Besides statistical-based context, as attested by modern grammar textbooks (Al-Rajihi, 1998 [1972]; Hassan, 2010), numbers tend to characterize and specify the nature of things. Proper nouns and past tense have the same significant loading; they share with numbers the non-abstract generation of meanings. Named after human, location, organization, software and hardware machines, equippeople, scientific terminology, temporal concepts, and events (Al-shalabi et al., 2011), proper nouns are not context-dependent, but ubiquitous and intertextually marked. As numbers can function autonomously in texts, proper nouns can stand as independent referents. Hassan (2010) argues that proper nouns support the concrete description of the text, conveying self-fulfilling rather than imaginative meanings. Past tense, in turn, is used to express past states, specific, and completed actions (Alhawary, 2016; Sekhri, 2008). The past records history, relate stories, deliver news, and document completed events. It is hard to find texts that use the present and

past tenses at the same time. While present state is sometimes abstract since it does not specify beginning-or-ending time, past forms are often accompanied by words or phrases that signify a specific point in time. In a wider context, best structures that do this job are adverbs of time, by answering the question 'when'? (Qabash, 2002; Ryding, 2005). Specificity is expressed by time adverbials. Relevant to the same function, *fī* preposition, referring to the place and time (Sibawayhi, 1988 [1975]), and commonly known as a locative and temporal marker (Wright, 1996; Mehdi & Jaradat, 2021; Zeddari, 2016), strengthens the concrete signification.

Overall, abstract relations, continuity of action, and factual expression of ideas are explored in the positive part. On the negative side, features tend to express statistical knowledge, specifying the nature of things without reference to abstract idealization. Negative co-occurrence also conveys self-fulfilling meanings, timed and localized, and concrete description of the text. Factor 4 is thus labeled dimension of *Abstract factual vs. Concrete factual style*. Figure 6 graphs the scale of dimension 4.

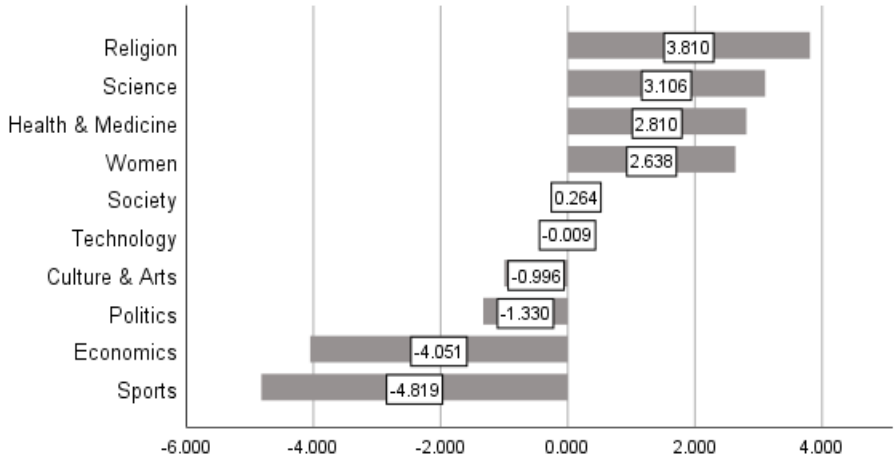

Figure 6. Dimension of *Abstract factual vs. Concrete factual style*

There are four subgenres of MA located noticeably on the positive line of *Abstract vs. Concrete style* dimension. To put them in a respective way, they are *Religion*, *Science*, *Health & Medicine*, and *Women*. *Sports*, *Economics*, *Politics*, and *Culture & Arts* are respectively negative subgenres; that is, they embody the concrete factual style of this dimension. The rest, *Society* and *Technology*, have an in-between position, with a value on the scale close to 0.0. *Economics* is a highly representative example of concrete factual style.

4.5 Interpretation of Factor 5

Factor 5 includes 3 positive features and 5 negative features (Table 3). To start with the positive group, Hassan (2010) ascribes numerous functions to the definite nouns, such as

definition, identification, distinction, descriptions, removal of ambiguities. Still, such functions are interrelated. According to Alhawary (2016), definite nouns re-introduce already mentioned rather than new information, and shared knowledge among writers and readers; they, now to Ryding (2005), specify the meaning of a noun, supporting generic use of language. A following co-occuring feature is adjectives. As pointed out by Biber (1988), adjectives are used in language to expand and elaborate the information presented in a text. Traditionally, adjectives are employed for continuation, contextualization, clarification, and, to some extent, certainty (Hassan, 2010).

The third lexico-grammatical feature exiting with an acceptable loading in Factor 5 is passive form. Biber (1988) states that the use of passive forms results in static, more abstract presentation of information. Universally across languages, passives hide the agent and focus on the patient of the action. In a particular context, they are said to be employed for thematic purposes (Weiner & Labov, 1983). Yet broadly, there a number of reasons accounting for the use of passive forms. In a textual analysis, Khalil (1988) succinctly states four major reasons that prompt the writer to conceal the identity of the agent: (a) fear of the agent who has done something wrong, (b) wish to protect or not to embarrass the agent who has done something wrong, (c) aversion to the agent who has been involved in an abominable, (d) the passive is also used in contexts where the focus is on prohibitions. All these reasons affirm non-concrete, indirect, subjectless demonstration of ideas, distinct qualities of non-narrative style of writing. The narrative style, by contrast, involves the presence of the speaker, either as a personage reported in the third-person noun or first-person narrator.

Therefore, shared knowledge, general meanings, and non-narrative description of ideas are perpetuated by the use of definite nouns, adjectives, and passives. The negative group functions in an opposite direction. The content, as is will be seen, is narrative.

As stated above, forms of first-person pronouns designate the presence of the speaker. Applicable to that, the writer can be the narrator of the event, the teller of the story, or the eyewitness deliverer of the news. Subject-based language reflects narrative stance where the doer of the action, or the one who is supposed to participate in discursively, is bluntly unveiled. Negation particles, in turn, can be deemed indicators of narrative content. The narrator does not stop at informing the addressee, conveying positively-driven ideas, but goes on to relate the story and stretch the discourse with a bit of elaboration, illustration, and sometimes interpretation. They are used with both nominal and verbal sentences. The use of negation requires respectable grasp of language, because its use does not only affect the morphological formation of following verbs and nouns but also the syntactic roles of words in the clause formation. *laysa* particle, for instance, may occur in two syntactic structures, stripping or negative contrast constructions. In the first one, it deletes everything in a clause under identity with corresponding parts of a clause. In the second structure, it is implemented as an indeclinable negative particle, stronger than *lā*, to deny the sentence part merged with.

Hussein (2011) summarizes an immense body of research on the lexico-grammatical expression of reasoning in Arabic. Regarding *kay*, a future particle co-occuring with other features in the negative group, he notices that the particle is employed in purposeful

reasoning, the presupposition of a purpose realized in the future. This function, asserted somehow concisely in classical research (Al-Malqi, 1975; Ibn Hicham, 1985; Sibawayhi, 1988 [1975]), is a marker of powerful motive for the persuasive language found in narrative accounts.

Conditional clauses, as Sartori (2019) and Hammadi (2019) unanimously explain, stand in the opposite side with real, factual speech. In his syntactic-semantic study, Alfraidi (2017) assumes that conditionals convey modality meanings and time references. In general, conditionals assert the absence of conditioned actions at the time of speaking/writing. To call all the three features, conditionals, *lākinna* particle, and future with *kay* particle share what might be called 'wishful' use, the expression of wish to have an action realized in the future, but it is deprived because of concession, strict conditions, a long journey that is needed, or there is not a realistic expectation that it will actually occur.

Negative features are indicators of a narrative content, detailed, first-person-dominated, and distinguished by stylistic structures, the syntactic constructions where *kay* and conditionals play no small part. Put everything together, Factor 5 is labeled *Generic non-narrative use of language vs. Narrative content* (Figure 7).

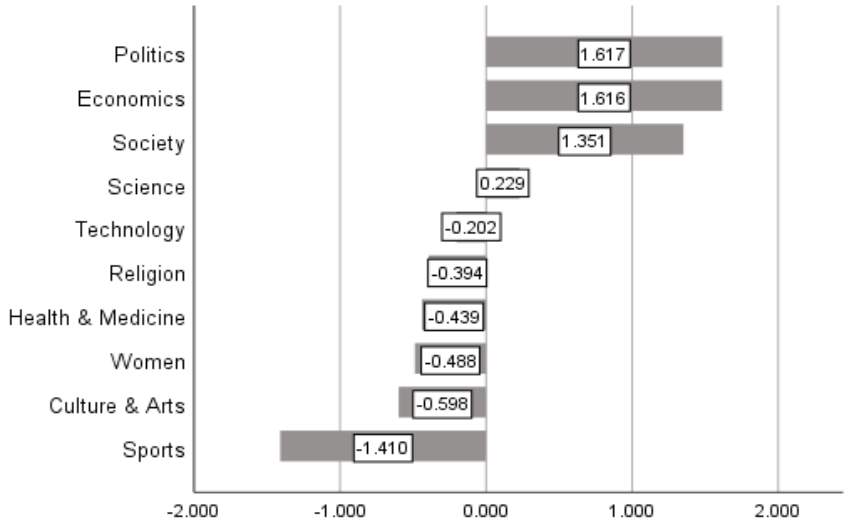

Figure 7. Dimension of *Generic non-narrative use of language vs. Narrative content*

Politics, *Economics*, and *Society* are highly positive on dimension 5. *Science* is also positive. Compared to the former subgenres, its position is insignificant. The rest of subgenres are posited on the negative line, indicating *narrative content*. They are categorized in such a descending way, from highly to minimally negative, as follows: *Sports*, *Culture & Arts*, *Women*, *Technology*, *Health & Medicine*, and *Religion*.

4.6 Interpretation of Factor 6

Factor 6 includes 6 positive features and 7 negative features (Table 3). The use of *ʾaṣbaḥa*,

demonstrates a situational change, a noticeable shift from one particular state to another. Accordingly, it is used to deliver new information unknown to the reader, a state of affairs, dynamic context, or newly acquired qualities regarding a person, topic, phenomenon, etc. The change encouraged by the use of *ʾaṣbaḥa* is most likely suspended and emphasized, regardless of remarkable events that happended or might happen in the future. Stable change is also exemplified by the use of *kāna* and *mā zāla* (Al-Matlabi, 1986), a fairly significant feature co-occurring with *ʾaṣbaḥa*.

Future with *sawfa* particle conveys meanings of predictability and insistence (Abunowara, 2005). Different from future with *sa* particle, the partice *sawfa* is known among grammarians as a remote future marker (Al-Matlabi, 1986; Hassan, 2010; Sibawayhi, 1988 [1975]). The expression of future action, in this vein, bears lesser degree of certainty than *sa* particle. The use of *sawfa* suspends the present status for a period indefinite in time. Relevant to the action suspension function, lack of certainly is expressed by verbs of probability and *lākinna* particle, too. In fact, future with *sawfa* particle and verbs of probability converges at a clear point, characterized by abstract and subjective estimation. It is this subjective expression which proves the presence of the first speaker in texts, namely the extensive use of first-person pronouns. By a concession style (Hassan, 2010), *lākinna* particle is used, in tunr, to sustain stability.

The negative features are opposite to this functional line; the lexico-grammatical expression is rather progressive, specifically in place. By answering the question ‘where?’ (Qabash, 2002), place adverbials indicate the location of the action. Considered as locative nouns and pronouns, they express spatial, existential, or directional knowledge (Ryding, 2005).

Although *ṣāra* denotes the sense of *ʾaṣbaḥa* (Hardie & Ibrahim, 2021), co-occurring with the adverbs of place and exceptive sentences, and other negative features, it signifies a different meaning. Derived from the noun *ṣayrūra*, process, the actions or states that follow *ṣāra* refers to one of a set of stages involved in the completion of a particular event. It is thus motion-related rather that stability-specific item. Its sisters *bāta* and *ḍhalla* are presumed to play the same role of movement marking. In other words, though indicating the persistence of a particular action in a certain state of shape or occurrence, the state that is persisted is related primarily to particular spatial conditions, revealed by the writers in either abstract or concrete description; these conditions make up the derivational nouns that can be adopted from such verbs as *bāta* and *ḍhalla* as corresponding to *bayt* and *ḍhill*, house and shadow. Setting the conditions aside, the actions are susceptible to change.

Moreover, the use of *hasaba* and third-person pronouns preserves the development of location-based description of events. *hasaba*, according to Ibn Mandhour (1968), is a word class of probability, yet the probability of *hasaba* designates that it is pragmatically relative to a following reference. The major function of is to cite sources, report statistics, or quote testimonials, hence a marker of reference. In much the same way, future with *ʾalā* or *ʾamā* particle reinforces the developmental, interactive flow of information (Ibn Hisham, 1985; Sibawayhi, 1988 [1975]), phrasing interrogatives whose *jawāb* takes place in the future. Depending on this assessment of the two clustered groups, positive and negative features,

Factor 6 is interpreted as dimension of *Stable suspension vs. Locative development* (Figure 8).

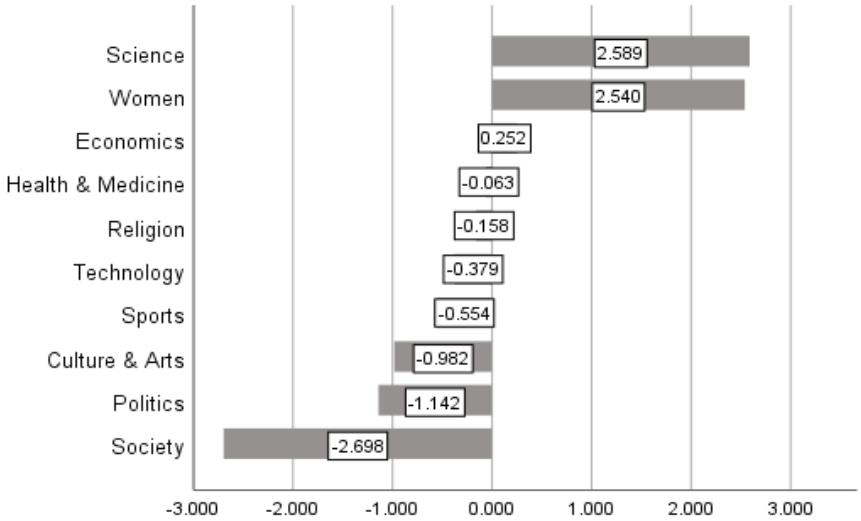

Figure 8. Dimension of *Stable suspension vs. Locative development*

The MA subgenres that contain *Locative development* are, respectively stated, *Politics*, *Society*, *Sports*, *Economics*, *Health & Medicine*, and *Culture & Arts*. The other subgenres are representations of stable content, especially *Science*, drawing a remarkable positive leap.

5. Recapitulations and Conclusions

The present corpus-based paper explores six dimensions of variation in media Arabic genre. The six-dimensional design, an interpreted output of EFA, documents the lexico-grammatical similarities and differences among ten subgenres of MA: *Politics*, *Economics*, *Society*, *Sports*, *Science*, *Culture & Arts*, *Technology*, *Women*, *Health & Medicine*, and *Religion*. In an expository and explanatory way, the study is concisely demonstrated in its total methodological application of Multi-dimensional framework, briefly describing the compilation of a representative MA corpus, the design of the coding lexico-grammatical scheme, strategies of annotation, techniques and steps of analysis. The interpretation of factors results in six dimensions of variation in media Arabic genre: (1) *Situation-dependent reference*, (2) *Open dynamic expression of assertion*, (3) *Interactive writer-reader interdependency discourse*, (4) *Abstract factual vs. Concrete factual style*, (5) *Generic non-narrative use of language vs. Narrative content*, and (6) *Stable suspension vs. Locative development*. Based on the functional co-occurrence of lexico-grammatical features, subgenres are found distinct or relatively marked unidimensionally, bidimensionally, or multidimensionally, revealing a great amount of linguistic similarities and differences within the internal ecology of media Arabic.

The study proves that there is a significant multi-dimensional co-occurring structure, analyzed in an exploratory fashion, and interpreted on a functional basis. Granted that MD is equipped with rigorous exploratory and interpretive adequacy, it holds unquestionable value

to the scientific inquiry of language. Yet to put the knowledge into practice, despite the research journey, designated by questions and objectives, culminates in completion, the study as well as the model it adopts alone do not suffice. However, raising one of the substantial questions left as vividly critical Chomskyan legacy is never vain: How is the knowledge of language put to use? Put practically, what can we *do* with what we *know* about media Arabic?

Even if the present endeavor has been much of theoretical and exploratory importance, what lies beyond the scope of it is more fruitful exploitation of its analyzed, interpreted, and discussed results. In fact, answers to the above question are many, but language pedagogy, the disciplined understanding of how a subject can be effectively learned, gets much responsive to the front. To give an example from TAFL, syllabus designers and teachers of Arabic can take advantage of the drawn conclusions to know what parts of language content to include, illustrate, elaborate, and assess at a particular, appropriate proficiency level. Instructors of Arabic for specific purposes, interested in a particular subgenre such as *Politics*, or a dimension such as *Abstract factual vs. Concrete factual style*, also can be enlightened with applied knowledge. Implications of the study for broader research relate, in turn, to a cross-linguistic quest, best expressed by Biber (2014: xxxiv) as the need to confirm *the universal parameters of language variation*. This latter pursuit is still too early, at least in respect to languages which have not been corpus-linguistically investigated on an extensive scale. Within the Arabic linguistic community, the study encourages projects macroscopic in Multi-dimensional application that requires a well coordinated, specialized crew in Arabic theoretical linguistics, corpus design, and computer science. Research can further implement confirmatory approaches to examine the dimensions of variation explored by previous studies. In another relevant context, since the present study has been limited to having a dimensional insight into the linguistic fabric of the media Arabic genre, it urges an active qualitative discussion about the motives and causes that account for subgenres' specific positions on particular dimensions. The ultimate end, at last, in exploring part of Arabic *terra incognita*, as it were, through attested language data (corpus) is to reduce the internal ambiguity on media Arabic by grasping novel empirical information about its genre and subgenres, bearing in mind that questioning the nature of knowledge of language, and its explanatory power, must not be preceded by any possible practical actions, lest there are nothing to justify a given practice but mere personal experience and unguided intuition.

References

Abbas, A. F. (2009). *Asālibu al-bayān*. Amman, Jordan: Dār an-Nafā'iss.

Abdelali, A., Darwish, K., Durrani, N., & Mubarak, H. (2016). Farasa: A fast and furious segmenter for Arabic. *NAACL*-2016.

Abdul Razak, R. Z. (2011). *Modern media Arabic: A study of word frequency in world affairs and sports sections in Arabic newspapers* (Unpublished doctoral dissertation). University of Birmingham, Birmingham, England.

Abunowara, A. (2005). Sa- and sawfa as modal forms: New evidence from informant testing. *International Journal of Arabic-English Studies*, *6*, 85-92. https://doi.org/10.33806/ijaes2000.6.1.7

Adham, M. (n.d.). *Al-maqālu aṣ-ṣiḥafī*. Cairo, Egypt: Maktabat al-Angelo al-Miṣriyya.

Al-Bakkae, M. (2010). *Al-'i'elāmu wa-l-lugha*. Damascus, Syria: Dār Nīnwa.

Al-Dobaian, A. S. (2018). A syntactic analysis of Arabic tense and aspect. *Advances in Languages and Literary Studies*, *9*(6), 82-92. https://doi.org/10.7575/aiac.alls.v.9n.6p.82

Al-Farahidi, A. (1998). *Al-'ayn*. Beirut, Lebanon: Dār al-Fikr al-'Arabī.

Alfraidi, T. R. (2017). *Conditional sentences in Modern Written Arabic* (Unpublished doctoral dissertation). University of Exeter, Exeter, England.

Al-Hamad, A., & Al-Zuebi, Y. (1993). *Al-mo'jamu al-wāfi fī 'adawāti an-naḥwi al-'arabī*. Giza, Egypt: Dār al-'Amal.

Alhawary, M. T. (2016). *Arabic grammar in context*. New York, NY: Routledge.

Al-Horais, N. (2017). On negation and focus in Standard Arabic: Interface-based approach. *Journal of Universal Language*, *18*(1), 1-34. https://doi.org/10.22425/jul.2017.18.1.1

Al-Jurjani, A. (1982). *Al- muqtaṣidu fī sharḥi al-'īḍāḥ* (B. K. Al-Marjane, *Tahq.*). Damascus, Syria: Dār ar-Rashīd.

Al-Mallakh, M. (2009). *Az-zmānu fī l-lugha: Binyātuhu at-tarkībiyatu wa dalāliya.* Rabat, Morocco: Dār al-'Amān.

Al-Malqi, A. A. (1975). *Raṣfu al-mabāni fī sharḥi hurūfi al-ma'ānī* (M. A. Al-Khattat, *Tahq.*). Damascus, Syria: Arabic Language Academy.

Al-Matlabi, M. Y. (1986). *Az-zamānu wa l-lugha*. Cairo, Egypt: Al-Hayāt al-Miṣriya li al-Kitāb.

Al-Mubarrad, A. A. M. (1967). *Al-Muqtaḍab* (A. K. M. Adima, *Tahq.*). Cairo, Egypt: Higher Councel for Islamic Affairs.

Al-Muradi, H. B. Q. (1992). *Al-janyu ad-danī fī hurūfi al-mabāni*. Beirut, Lebanon: Dār al-Kutub al-I'lmiya.

Al-Rajihi, A. (1998 [1972]). *At-tatbīqu an-nahwī*. Alexandria, Egypt: Al-Ma'rifatu al-Jāmi'iya.

Al-Shahri, B. A. (2004). *'Istratījiyātu al-khitābi: Muqārabatun lughawiyatun tadāwuliya*. Beirut, Lebanon: Dār al-Kitāb al-Jadīd al-Muttaḥida.

Al-Shalabi, R., Kanaan, G., Al-sarayreh, B., & Khanfar, K. (2009). Proper noun extracting algorithm for Arabic language. *International Journal of Computer, the Internet and Management*, *19*(1), 45-53.

Al-Shamsan, A. I. (1987). *Hurūfu al-jar: Dalalātuha wa-'alāqātuha*. Riyadh, Saudi Arabia: Malek Saud University.

Al-Sirafi, A. S. (2008). *Sharḥu kitābi Sibawayhi* (H. A. Mahdali & S. A. Ali, *Tahq.*). Beirut, Lebanon: Dār al-Kutub al-I'lmiya.

Al-Suyoti, J. A. (2006). *Ham'u al-hawāmi'e fī sharḥi jam'i al-jawāmi'e*. Beirut, Lebanon:

Dār al-Kutub al-I‘lmiya.

Al-Tamimi, M. A. (2016). Personal deixis in nahju al-balāgha. *College of Humanities Magazine, 47*, 317-348.

Al-Zamakhshari, A. (1974). *Al-mufaṣṣal fī ‘ilm al-‘arabiya* (M.S. Akil, *Tahq.*). Beirut, Lebanon: Dār al-Jīl.

Al-Zobaidi, M. M. H. (1971). *Tāju al-‘arūsi min jawāhiri al-qāmūs* (A. A. S. Farraj, *Tahq.*). Kuweit, Kuweit: Kuweit Goverment Press.

Anis, I. (1994). *Min aṣrāri al-lugha*. Cairo, Egypt: Maktabat al-’Anjelo al-Miṣriya.

Atkins, S., Clear, J., & Ostler, N. (1992). Corpus design criteria. *Journal of Literary and Linguistic Computing, 7*(1), 1-16. https://doi.org/10.1093/llc/7.1.1

Basharat, A. M. (2017). ṣiyaghu al-ʾamri fī al-ʾiʿrābi ʾaw al-bināʾ: qirāʾatun jadīdatun fī al-binyati at-tarkībiyati wa-l-waḍīfiyati ad-dalāliya. *Majallat ʾItiḥādi al-Jāmiʿāti al-ʿArabiya, 14*(1), 111-129.

Beaman, K. (1984). Coordination and subordination revisited: Syntactic complexity in spoken and written narrative discourse. In T. Tannen (Ed.), *Coherence in spoken and written discourse* (pp. 45-80). Norwood, NJ: Ablex.

Belaid, S. (2006). Anqidhu al-lughata al-‘arabiyata mina aṣ-ṣiḥāfiyīn. In Higher Council for Arabic Language in Algeria (Ed.), *Munāfaḥātun fī al-Lughati al-‘Arabiya*. Tizi Ouzou, Algeria: Dār al-’Amal.

Belnap, K. (1986). *Complementation in Modern Standard Arabic: A corpus-based approach* (Unpublished Master’s thesis), Brigham Young University, Quebec, Canada.

Biber, D. (1986a). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*, 384-414. https://doi.org/10.2307/414678

Biber, D. (1986b). On the investigation of spoken/written differences. *Studia Linguistica, 40*, 1-38. https://doi.org/10.1111/j.1467-9582.1986.tb00759.x

Biber, D. (1987). A textual comparison of British and American writing. *American Speech, 62*(2), 99-119. https://doi.org/10.2307/455273

Biber, D. (1988). *Variation across speech and writing*. Cambridge, UK: Cambridge University Press.

Biber, D. (1989). A typology of English texts. *Linguistics*, *27*(1), 3-43. https://doi.org/10.1515/ling.1989.27.1.3

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing, 8*, 243-257. https://doi.org/10.1093/llc/8.4.243

Biber, D. (2014). Multidimensional analysis: A personal history. In T. B. Sardinha, & M. V. Pinto (Eds.), *Multi-dimensional analysis, 25 years on: A tribute to Douglass Biber* (Opening essay). Amesterdam, Netherlands: John Benjamins Publishing Company.

Biber, D., & Conrad, S. (2009). *Register, genre, and style*. New York, NY: Cambridge

University Press.

Biber, D., & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes, 11*(1), 1-34. https://doi.org/10.1080/01638538809544689

Brezina, V. (2018). *Statistics in corpus linguistics*. Cambridge, UK: Cambridge University Press.

Chejne, A. (1969). *The Arabic language: Its role in history.* Minneapolis, MN: University of Minneapolis Press.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(7), 1-9. https://doi.org/10.7275/jyj1-4868

Darwin. (1959). *On the origins of species*. New York, NY: D. Appleton and Company.

Dendane, Z., & Dendane, A. (2012). The negation in Arabic: An issue for translation. *Revue Maghr ḍbine des Langues*, *1*, 1-17.

Eid, M. (1983). On the communicative function of subject pronouns in Arabic. *Journal of Linguistics*, *19*(2), 287-303. Retrieved from https://www.jstor.org/stable/4175694

Friginal, E., & Hardy, A. J. (2014). Conducting Multi-dimensional analysis using SPSS. In T. B. Sardinha, & M. V. Pinto (Eds.), *Multi-dimensional analysis, 25 years on: A tribute to Douglass Biber* (pp. 297-321). Amesterdam, Netherlands: John Benjamins Publishing Company.

Gorsuch, L. R. (2015). *Factor analysis*. New York, NY: Routledge.

Grabe, W., & Biber, D. (1987). Freshman student writing and the contrastive rhetoric hypothesis. Paper presented at *SLRF 7*, University of Southern California, US.

Halliday, M. A. K., & Hasan, R. (1976). *Coherence in English*. London, UK: Routledge.

Halliday, M. A. K., & Hasan, R. (1985). *Language context and text: Aspects of language in a social semiotic perspective.* Oxford, UK: Oxford University Press.

Hammadi, S. S. (2019). Arabic and English conditional clauses: A comparative study. *The Eurasia Proceedings of Educational & Social Sciences*, *13*, 109-114.

Hammou, N. (2010). Al-ʻudūl an-nahwī fī al-lughati al-ʻarabiya: Jarīdatu as-shurūq al-yawmī namūdhajān. In Higher Council for Arabic Language in Algeria (Ed.), *Al-lughah al-ʻarabiya fī as-sihāfa al-maktūba* (pp.228-268). Alger, Algeria: Higher Council for Arabic Language.

Hardie, A., & Ibrahim, W. (2021). Exploring and categorising the Arabic copula and auxiliary kāna through enhanced part-of-speech tagging. *Corpora*, *16*, 305-335. https://doi.org/10.3366/cor.2021.0225

Hassan, A. (2010). *An-naḥwu al-wāfī.* Cairo, Egypt: Al-Maʻārif.

Holes, C. (2004). *Modern Arabic: Structures, functions, and varieties*. Washington, DC: Georgetown University Press.

Horn, R., & Martin, W. (2008). *EPS 624/725: Exploratory factor analysis.* [Class Tutorial/Handout]. Northern Arizona University, Flagstaff, AZ.

Hosni, M. M. (1997). *An- naḥwu as-shāfī*. Beirut, Lebanon: Ar-Risāla.

Hussein, O. I. (2011). Hurūfu at-ta'elīli ʿinda an-naḥwiyīna wa al-mufassirīn. *Majalat al-Ustadh, 159,* 2-60.

Ibn Aqil, B. A. A. (1997). *Sharḥu alfiyati Ibn Mālik.* Beirut, Lebanon: Dār al-Jabal.

Ibn Faris. (1971). *Maqāyīsu al-lugha* (A. Haroun, *Tahq.*). Damascus, Syria: Dār al-Fikr.

Ibn Hicham, J. (1985). *Mughni al-labībi 'an kutubi al-'a'ārib* (M. Al-Mubarak & A. M. Hamdallah, *Tahq.*). Damascus, Syria: Dār al-Fikr.

Ibn Malek, B. A. M. (1967). *Tashīlu al-fawā'idi wa takmīlu al-maqāsid* (K. M. Barakat, *Tahq.*). Cairo, Egypt: Dār al-Kitāb al-'Arabī.

Ibn Mandhour. (1968). *Lisānu al-'arab*. Beirut, Syria: Sadir.

Ibn Yaish, M. (2001). *Sharḥu al-mufaṣṣal*. Beirut, Lebanon: Dār al-Kutub al-'Ilmiya.

Igaab, Z. K., & Tarrad, I. R. (2019). Pronouns in English and Arabic: A contrastive study. *English Language and Literature Studies, 9*(1), 53-69. http://doi.org/10.5539/ells.v9n1p53

Khalil, A. (1988). The passive voice in English and Classical Arabic: Formation, type and function. *Bethlehem University Journal, 7,* 7-38. Retrieved from https://www.jstor.org/stable/26445057

Lee, D. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning and Technology, 5*(3), 37-72. Retrieved from https://ro.uow.edu.au/artspapers/598

Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.), *Directions in corpus linguistics* (pp. 105-127). Berlin, Germany: De Gruyter Mouton.

Leech, G. (2007). New resources, or just better old ones?. In M. Hundt, N. Nesselhauf, & C. Biewer (Eds.), *Corpus linguistics and the web* (pp. 134-49). Amsterdam, Netherlands: Rodopi.

Marckworth, M. L., & Baker, W. J. (1974). A discriminant function analysis of co-variation of a number of syntactic devices in five prose genres. *American Journal of Computational Linguistics, Microfiche, 11.*

Martin, J. (2001). Language, register and genre. In A. Burns, & C. Coffin (Eds.), *Analysing English in a global context: A reader* (pp. 149-166). London, UK: Routledge.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Method, theory, and practice.* New York, NY: Cambridge University Press.

Mehdi, M. F., & Jaradat, M. (2021). Adverbs of time in Arabic and English: Comparative study. *International Journal of Linguistics, 13*(1). 85-97. http://doi.org/10.5296/ijl.v13i1.18254

Mohamed, G., & Hardie, A. (2019). Approaching text typology through cluster analysis in Arabic. In T. McEnery, A. Hardie & N. Younis (Eds.), *Arabic corpus linguistics* (pp. 201-228). Edinburgh, Scotland: Edinburgh University Press.

Mouawad, S. (2008). *Ḥurūfu al-ma'ānī*. Tripolis, Lebanon: Al-Mu'assasa Al-Ḥadītha li al-Kitāb.

Mulyadi, D. (2021). A study on relative clause in English, Indonesian and Arabic languages: A comparative analysis. *Journal of Language Teaching and Learning, Linguistics and Literature*, *9*(1), 109-115. http://doi.org/10.24256/ideas.v9i1.1804

Naknawa, D. (2020). Adjectives as word class in Arabic: A linguistic analysis. *Journal of Letters*, *49*(1), 92-105.

Nunan, D. (2008). Exploring genre and register in contemporary English. *English Today, 24*(2), 56-61. https://doi.org/10.1017/S0266078408000217

Ochs, E. (1979). Planned and unplanned discourse. In T. Givon (Ed.), *Discourse and syntax* (pp. 51-80). New York, NY: Academic Press.

Ostler, N. (2006). *Empires of the word: A language history of the world*. New York, NY: Harper Perennial.

Ostrabadhi, R. (n.d.). *Charḥu kāfiyati ibn al-Ḥajib*. Beirut, Lebanon: Dār al-Kutub al-'lmiya.

Paltridge, B. (1996). Genre, text type, and, and the language classroom. *ELT Journal*, *50*(3), 237-243. https://doi.org/10.1093/elt/50.3.237

Qabash, A. (2002). *Al-kāmilu fī an-naḥwi wa-ṣ-ṣarfi wa-al-'i'rāb*. Beirut, Lebanon: Dār al-Jīl.

Quirk, R., & Greenbaum, S. (1973). *A university grammar of English*. Oxford, UK: Oxford University Press.

Ryding, K. C. (2005). *A reference grammar of Modern Arabic*. Cambridge, MA: Cambridge University Press.

Saleh, H. A. (2007). Al-lughatu al-'arabiyyatu bayna al-mushāfahati wa-at-taḥrīr. *Buhūth wa-Dirāsāt fī al-Lisāniyyāt al-'Arabiyya*, *66*, 72-81.

Salim, A. A., & Ahmad, M. A. (2015). Semantic structure of relative clauses in English and Arabic: A contrastive study in Sudanese context. *European Academic Research, 3*(8), 9307-9345.

Sartori, M. (2019). A relational approach to modern literary Arabic conditional clauses. In T. McEnery, A. Hardie, & N. Younis (Eds.), *Arabic corpus linguistics* (pp. 143-169). Edinburgh, Scotland: Edinburgh University Press.

Sekhri, O. (2008). *Problems in translating tenses from English into Arabic: The present perfect: A case study* (Unpublished doctoral dissertation). University of Constantine, Constantine, Algeria.

Sharaf, A. (1991). *Al-lughatu al-'i'lāmiya*. Beirut, Lebanon: Dār al-Jil.

Sibawayhi, A. (1988 [1975]). *Al-kitāḅ* (A. Haroun, *Tahq.*). Cairo, Egypt: Al-Khanji.

Sinclair, J., & Ball, J. (1996). Preliminary recommendations on text typology. *EAGLES (Expert Advisory Group on Language Engineering Standards)*. Retrieved from http://www.ilc.cnr.it/EAGLES96/texttyp/texttyp.html

Steen, G. (1999). Genres of discourse and the definition of literature. *Discourse Processes, 28*, 109-120. https://doi.org/10.1080/01638539909545075

Swales, J. (1990). *Genre analysis*. Cambridge, MA: Cambridge University Press.

Weiner, J. E., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics, 19*, 29-58.

Winter, E. (1982). *Towards a contextual grammar of English: the clause and its place in the definition of sentence.* London, UK: George Allen and Unwin.

Wright, W. (1967). *A grammar of the Arabic language*. Cambridge, MA: Cambridge University Press.

Zaydan, J. (1988). *Al-lughatu al-'arabiyatu kā'inun ḥay*. Beirut, Lebanon: Dār al-Jīl.

Zeddari, I. (2016). A comparative analysis of locative and directional motion events in English and Arabic. *International Journal of Linguistics*, *8*(4), 133-149. https://doi.org/10.5296/ijl.v8i4.9725

Zoubae, A. A. (2016). Al-binyātu aṣ-ṣawtiyatu wa-dalāliyatu li-ḥarfi al-kāf: Dirāsatun muqārinatun bayna al-lughatayni al-'arabiyati wa-al-'ibriya. *Majallat Markaz Babel li-Dirāsāti al-Lisāniya*, *6*(1), 104-130.

Copyrights