

**WRITTEN MOROCCAN ARABIC ON THE INTERNET:
TYPES OF VARIATION AND LEVELS OF REGULARITY**

**ÁRABE MARROQUÍ ESCRITO EN INTERNET:
TIPOS DE VARIACIÓN Y NIVELES DE REGULARIDAD**

العربية المغربية الدارجة المكتوبة في الإنترنت: أنواع البدائل ودرجات الانتظام

Marcin Michalski*
Universidad Adam Mickiewicz en Poznań

Recibido: 12/04/2025

Aceptado: 22/10/2025

BIBLID [1133-8571] 32 (2025) 5-30

Abstract: For some time, Moroccan Arabic has been more and more extensively used in writing in various areas of the Internet: whether by ordinary Internet users, more or less amateur authors of expository texts aimed at the general public (the Wikipedia) or by journalists. Since Written Moroccan Arabic has no orthography, i.e. a generally accepted spelling standard, variation affecting a considerable number of forms is one its conspicuous features.

The aim of the present study is to discuss certain graphic phenomena related to writing Moroccan Arabic in Arabic script manifesting themselves in three text genres: readers' comments (on a major Moroccan news platform), articles in the Moroccan Wikipedia, and journalistic texts (in a popular online newspaper). To this purpose, a theoretical and methodological apparatus is proposed which includes, beside other notions, the concept of variation, understood as a relation between graphic words (variants) which have the same meaning and pronunciation but different graphies. Three basic types of variation are distinguished: (i) qualitative (the difference between the variants consists in each of them containing a different graph – or different graphs – in the same position), (ii) quantitative (the differentiating element is a graph present in one form and absent from the other), and (iii) linear (the differentiating element is a space).

The results of the study show that the graphy of texts written in Moroccan Arabic on the Internet differs in certain respects from that of printed literary texts, although some shared features can also be identified. The analysis also reveals that the three genres under examination differ from each other in terms of the occurrence of certain graphic phenomena, even though sometimes the differences are more a matter of frequency than clear-cut division. Finally, the levels of stability, or regularity, of particular genres vary, sometimes contrary to initial expectations.

Key words: Moroccan Arabic, Written Moroccan Arabic, Written Darija, Arabic graphy, Arabic script, Internet Arabic

Resumen: Desde hace algún tiempo, el árabe marroquí se utiliza cada vez más por escrito en diversas áreas de Internet: tanto por los usuarios comunes o aficionados que crean varios textos expositivos dirigidos al público general (como en Wikipedia), como por periodistas. Dado que el árabe marroquí escrito no cuenta con una ortografía, es decir, una norma de escritura generalmente aceptada, una de sus características más notorias es la variación que afecta a un número considerable de formas.

* Email: mmich@amu.edu.pl. ORCID: <https://orcid.org/0000-0002-4550-905X>

El objetivo del presente estudio es analizar ciertos fenómenos gráficos relacionados con la escritura del árabe marroquí en alfabeto árabe que se manifiestan en los tres géneros textuales siguientes: comentarios de lectores (en una importante plataforma de noticias marroquí), artículos en la versión marroquí de la Wikipedia y textos periodísticos (en un popular diario digital). Con este objetivo, se propone un aparato teórico y metodológico que incluye, entre otras nociones, el concepto de variación, entendida como la relación entre palabras gráficas (variantes) que tienen el mismo significado y pronunciación, pero diferente representación gráfica. Se distinguen tres tipos básicos de variación: (i) cualitativa (la diferencia entre las variantes consiste en que cada una contiene un grafema distinto – o grafemas distintos – en la misma posición), (ii) cuantitativa (el elemento diferenciador es un grafema presente en una forma y ausente en la otra), y (iii) lineal (el elemento diferenciador es un espacio).

Los resultados del estudio muestran que la representación gráfica de textos escritos en árabe marroquí en Internet difiere en ciertos aspectos de la de los textos literarios impresos, aunque también se pueden identificar características compartidas. El análisis revela, además, que los tres géneros examinados difieren entre sí en cuanto a la aparición de ciertos fenómenos gráficos, aunque en ocasiones las diferencias son de carácter estadístico y no absoluto. Se observa también que el nivel de regularidad varía según el género, a veces de manera contraria a lo que se podría esperar inicialmente.

Palabras-clave: Árabe marroquí, árabe marroquí escrito, dariya escrita, escritura árabe, alfabeto árabe, árabe en Internet.

المخلص: منذ فترة أصبحت العربية المغربية الدارجة تُستخدم بشكل متزايد في الكتابة في مختلف فضاءات الإنترنت، سواء من قبل المستخدمين العاديين، أو من قبل الهواة المؤلفين لنصوص وصفية موجهة إلى الجمهور العام (كما في ويكيبيديا)، أو من قبل الصحفيين. وبما أن الدارجة المغربية المكتوبة ليست لها قواعد إملائية معترف بها بشكل عام، فإن وجود البدائل الكتابية في عدد كبير من الحالات يُعدّ من أبرز سماتها. يناقش هذا البحث بعض الظواهر الشكلية المرتبطة بكتابة الدارجة المغربية بالحرف العربي كما تظهر في ثلاثة أنواع من النصوص: تعليقات الزوّار (على منصة إخبارية مغربية كبرى)، ومقالات في ويكيبيديا المغربية، ونصوص صحفية (في صحيفة إلكترونية شهيرة). يُقترح فيه إطار نظري ومنهجي يشمل، إلى جانب مفاهيم أخرى، مفهوم البديل الكتابي والذي يعني كلمة مكتوبة تشارك كلمة مكتوبة أخرى كلاً من المعنى والنطق وتختلف منها في شكلها الكتابي. هناك ثلاثة أنواع أساسية من البدائل الكتابية: (1) البديل النوعي (حيث يتم التمييز بين بديلين بفضل استخدام حرفين مختلفين في نفس الموضع في كل منهما)، (2) البديل الكمي (حيث يكون العنصر المميّز هو وجود حرف في بديل وغيابه في الآخر)، (3) البديل المعتمد على الفصل والوصل (العنصر المميّز هو وجود مسافة في بديل وغيابها في الآخر). تُظهر نتائج البحث أن كتابة النصوص بالدارجة المغربية على الإنترنت تختلف من حيث بعض الخصائص عن تلك التي تمارس في النصوص الأدبية المنشورة في صيغة الكتاب، مع وجود بعض السمات المشتركة. كما كشف الدراسة عن اختلاف الأنواع الثلاثة من النصوص فيما يتعلق بوجود بعض الظواهر الكتابية، ولو أن هذا الاختلاف يظهر بمثابة كثرة أو قلة استعمال صيغة معينة ولا وجودها أو غيابها بشكل مطلق. يلاحظ كذلك أن كل نوع من النصوص التي تمت دراستها يختلف من حيث درجة الانتظام، أحياناً على عكس التوقعات الأولية.

الكلمات المفتاحية: الدارجة المغربية، الدارجة المغربية المكتوبة، الدارجة المكتوبة، التمثيل الخطي بالعربية، الحرف العربي، العربية على الإنترنت.

1. Introduction

The term Written Moroccan Arabic (henceforth WMA) refers to a phenomenon that has been drawing increasing interest on the part of Arabists in the recent years. In the present study, it is understood, more specifically, as (particular varieties of) Moroccan Arabic (henceforth MA) used in writing, in contradistinction to speech, in Arabic script.⁽¹⁾ The recent surge in the use of WMA is owed, in part, to the change of attitudes towards MA, as opposed to Standard Arabic (henceforth SA), with the new generations viewing more favourably the use of the former in writing as a more natural mode of expression, with its promotion by some cultural figures and its recognition on the State level through Article 5 of the Moroccan Constitution of 2011 stipulating that the State acts to protect “the dialects and cultural expressions” used in Morocco. Partly, it has been facilitated by the emergence of new communication technologies such as the Internet with the social networking services, blogs, forums, etc. as well as various personal messaging systems.

(1) Systems used for writing varieties of Arabic based on Latin script, which do not concern us here, have various names, e.g. Arabizi, Arabish, Arabic chat alphabet. In the particular case of MA, usually referred to *darija* in English (from MA *d-dariža* ‘colloquial language’), the term “e-darija” is used by some linguists, especially Caubet (e.g. 2005, 2012, 2013, 2017, 2018). For a systematic outline of this system, see Moscoso García (2009).

That MA has been used in writing on an increasing scale, in both printed and digital texts, private and public, has been noted in many studies. For its use in private communication, see Berjaoui 2001; Benítez Fernández 2003; Caubet 2004, 2012: 388-399, 2013, 2017 and Moustauoui Srhir 2016: 113-115). The introduction of WMA into the public and semi-official sphere, including the media, both printed and electronic, has been discussed in Aguadé 2012, Benítez Fernández 2012a and b, Caubet 2012, Miller 2012, 2017, Elinson 2013: 717-719, Hoogland 2013b, 2018; Langone 2003, Moustauoui Srhir 2016: 110-113 and Pennisi 2020a and 2023.⁽²⁾

The graphy used in particular MA texts is to a great extent improvised, often chaotic and tends to be internally inconsistent. Linguistic aspects of WMA in Arabic script, in terms of regularities and variations, have been discussed by a rather limited number of scholars. Aguadé (2005) and (2013) are case studies of the use of WMA in particular books, while Aguadé (2006) provides a general picture of the WMA spelling practice based on a heterogeneous selection of texts. Observations on WMA graphy in various genres can be found in Moustauoui Srhir (2013: 119-131), while Hoogland (2013a) analyzes a corpus of texts belonging to various genres with the normative aspect in mind. Some additional graphic phenomena in WMA literary texts were described by Michalski (2016) and his larger study in (2019), with a focus on qualitative variation (see Section 2.2 below). Chapters or passages devoted to the graphy of MA are in Benítez Fernández (2010: 218-220). Caubet (2017: 133-136) describes some aspects of the MA graphy online, while Pennisi (2020b: 137-138) deals with the graphy used in Mġarfāwī et al. 2017, a dictionary of MA published in Morocco.

The present paper is hoped to show that the theoretical apparatus proposed in Michalski (2019) for the analysis of literary texts written in WMA can be used for description of WMA used on the Internet. Apart from the purely descriptive character of the study, another of its features has to be emphasized here: Rather than being a statistical approach, it should be considered as introductory study whose findings can be used in further, more complete, research based on statistical examination.

In the subsequent parts of this paper, after the necessary theoretical and methodological considerations in Section 2, the following topics will be addressed: How does the practice of writing MA in Arabic script on the Internet compare to its usage in printed literary texts (with Section 3 devoted to similarities and Section 4 to major differences)? Which types of variation affect which particular types of texts and which of these exhibit the highest degree of regularity (with these issues discussed in Section 5)?

2. Theoretical considerations

Most studies mentioned above are based on the representational conception, i.e. one whose point of departure are relations between sounds and graphs (letters, vocalization signs) which represent them. By contrast, the method used in the present paper is based on Michalski (2019: 50-82), an approach which in many respects favours the distinctivity conception, i.e. starting from relations between particular graphs, such as their mutual interchangeability or graphemic

(2) See Michalski 2016: 26 for studies dealing with modern literary creation in WMA, its use in contemporary *belles lettres*, including literary translations. WMA appears also in some dictionaries, grammars (e.g. al-Midlāwī (2019: 75-107), where he explicitly proposes “[t]he orthographic and phonetic rules of what he refers to as “Middle Moroccan Arabic”, educated Moroccan Arabic (apud Pennisi 2023: 299, footnote 8)) and textbooks and papers being proposals of its graphy but this issue exceeds the topic of the present study.

opposition (corresponding, at least to some extent, to free variation and phonological opposition, respectively, in phonology) and variation between graphic words. The representational function of graphs is of course referred to whenever necessary in the description since, it goes without saying, writing, although constituting a separate system, should not be analyzed in separation from spoken language, i.e. language signs it represents. The basic concepts are introduced in the following paragraphs.

2.1. Graphetic concepts

Any system of writing can be analyzed on two levels: graphetic and graphemic. Graphetics deals with graphs, i.e. units of written language that are distinguishable on a purely formal basis, without taking into consideration relations based on their meaning-differentiating functions. The factor of sounds being represented by particular graphs becomes relevant at some point, but as a rule, graphic words can be compared in terms of their meaning without taking it into consideration. Relations involving graphs and their meaning-differentiating functions are the object of graphemics (cf. Günther 1988: 64, 71). The relationship between graphetics and graphemics is, at least at some points, analogous to that between phonetics and phonemics.

The minimal and basic unit of description, whether on the graphetic or graphemic level, is *graph*. It corresponds to what is commonly referred to as letter (*ḥarf*) in Arabic studies, some clarifying comments are, however, required. First, it seems convenient to consider the following signs: ا, إ and ؤ as separate graphs rather than as combinations of one graph, i.e. the alif ا, with other graphs: the *hamza* and *waṣla*. The same holds true for و and ؤ as well for ع and ئ. The second issue is the status of vocalization signs: Should they be considered graphs even though they are hardly used in the Internet texts? Since they appear in some examples quoted from printed literary texts, they are listed in the Transliteration table below. Third, whether an Arabic graph is in the initial, medial or final position within a writing group, each of them entailing the use of a specific form, is fully predictable and of no further theoretical interest. Fourth, the individuation of graphs is not always an easy task. For instance, with the letter whose initial and word-medial forms are ك and ك, respectively,⁽³⁾ two final forms can be associated: ك and ك, e.g. for *fazag* ‘wet’, one can find, among other forms, the following: فازك⁽⁴⁾ and فازك⁽⁵⁾. Yet another possibility is the medial incomplete form used word-finally, as in ضرك⁽⁶⁾ *ḍræg* ‘cactus’. This suggests users’ uncertainty as to which form is “right” but some may even not be aware of the difference. This uncertainty is reflected in scholarly literature: Aguadé identifies the graphs as ك (2006: 259), while in Caubet it is ك (2017: 134). In the Moroccan Wikipedia, the form with ك is used as the title of the article ‘Honkong’⁽⁷⁾, but in the body of the article forms with ك are used. It seems that, perhaps for technical reasons, strict differentiating between these two graphs may be beyond control of ordinary users: both forms, هونكونك and هونكونك, look the same when pasted into the search box in the Wikipedia but yield different results: 4 pages for the former and 35 for the latter. Moreover, the count is misleading since the forms occurring on these 35 latter pages turn out to be هونكونك. The theoretical and methodological question which arises here is: Should the final ك and ك be considered two different graphs and receive two different transliteration symbols, or be described as two technically conditioned variants of one graph? In this study we adopt the latter, simpler solution and transliterate each form as <ḡ>.

(3) It is used in WMA to represent the sound [g] but this is irrelevant for the time being.

(4) E.g. in Fathī (2018a). In order not to inflate the volume of this paper, only examples of forms which are not very common are referenced.

(5) E.g. in Mzīyyān (2016).

(6) «Ḍræg n-nṣāra»...» (2023).

(7) «Honkong» (Wikipedia) (no date).

Another potentially difficult issue is that of **ث**, a graph with which two final forms: **ث** or **ث**, can be associated. This case is, however, clear since words with the final form **ث**, e.g. أرشيف *aršiv* ‘archive’, كييف *Kyiv* ‘Kyiv’, ستيف *Stiv* ‘Steve’, are used quite commonly, while the number of graphic words with final **ث** is so low that it can be ignored.⁽⁸⁾

Graphs and larger units which they constitute – i.e. graphic words, graphic sentences and graphic texts – can be transliterated into other graphic systems on the principle that one graph corresponds to one transliteration symbol and *vice versa*. The transliteration system used in the present study is given in Table 1 below.⁽⁹⁾ Transliteration is given in angle brackets < >. Space is transliterated as a low line: <_>. The system does not include the traditional Maghrebi letters **ف** marking *f* and **ق** marking *q* as they did not occur in the analyzed Internet text.

Graph	Transliteration	Graph	Transliteration	Graph	Transliteration
ا	<a>	ر	<r>	گ	<g>
أ	<ā>	ز	<z>	گ / گ	<g>
آ	<ā>	س	<s>	ل	<l>
إ	<a>	ش	<š>	م	<m>
ب		ص	<ṣ>	ن	<n>
پ	<p>	ض	<ḍ>	ه	<h>
ت	<t>	ط	<ṭ>	ة	<ḥ>
ث	<ṭ>	ظ	<ẓ>	و	<w>
ج	<g>	ع	<ʿ>	و	<w>
ح	<ḥ>	غ	<g>	ي	<ā>
خ	<x>	ف	<f>	ي	<y>
چ	<č>	ق	<q>	ئ	<y>
د	<d>	ف	<v>	ء	<o>
ذ	<d>	ك	<k>		

Table 1. The transliteration system of graphs used in WMA on the Internet

Vocalization signs are transliterated as superscript vowels reflecting their function in SA:

Graph	Transliteration
َ (<i>fatha</i>)	< ^a >
ِ (<i>kasra</i>)	< ⁱ >
ُ (<i>ḍamma</i>)	< ^u >
ْ (<i>šadda</i>)	< >
ْ (<i>sukūn</i>)	<°>

(8) Pace Aguadé (2005: 246), who gives **ث**, not **ث**, as the isolated form.

(9) The form of the transliteration symbols is, of course, not necessarily related to the sounds (phonemes) they are typically used to represent.

Table 2. *Transliteration of vocalization signs*

Graphs –as well graphic words, graphic sentences and graphic texts– which differ from others only with respect to features which are deemed irrelevant for their identification, i.e. size, colour, the kind of printing type (italic, bold), font or style, are said to be bound by *homography*. For instance, all graphic words in (1):

(1) دار – دار – دار (1)

are homographic and transliterated as <dar>. Graphs as well as larger graphic units that are not bound by homography are bound by the relation of *heterography*. Between units larger than graphs, three types of heterography can be distinguished: qualitative, quantitative and linear, depending on the nature of the distinctive unit(s).

Qualitative heterography relies on the quality – i.e. formal properties – of a graph. In other words, the distinctive unit is graph *x* in a given position of the graphic word as opposed to graph *y* in the same position, as in examples (2a) and (2b) (the distinctive units are transliterated in bold):

(2)

- a. دار <dar> — فار <far>
b. نقول <nqwl> — نڱول <nḡwl>

In *quantitative* heterography, the distinctive unit is the presence of a graph as opposed to its absence, as in examples (3a) and (3b):

(3)

- a. دار <dar> — دارو <darw>
b. لمغرب <lmḡrb> — لمغريب <lmḡryb>

Finally, in *linear* heterography, the distinctive unit is the space as opposed to its absence, as in examples (4a) and (4b):

(4)

- a. کانقرب <kanqrb> — کان قرب <kan_qrb>
b. فالددار <faldar> — ف الدار <f_alдар>

So far, no reference to meaning differentiated or sounds represented by the graphic units has been made (or this question was not decisive, as in the case of كَ or قَ). These functions come into play in the next section, devoted to graphemics.

2.2 Graphemic concepts

The graphic units exemplified in (2)-(4) are used to represent units of spoken (phonetic) language, which, in their turn, convey meanings. The identification of the relationship between a graphic words and a phonetic words its represents, as well as their meanings, can entail various theoretical and practical problems (see Section 2.4 for details). For the sake of convenience, it is assumed here that such identification can be carried out without difficulties for the examples used (even though no context is provided).

Heterography, whether qualitative, quantitative or linear, can be either distinctive or variational (non-distinctive). *Distinctive heterography* is a relation between heterographic

words that differ in meaning (and/or pronunciation). It is illustrated in examples (2a), (3a) and (4a) above, repeated here with the transcription of words assumed to correspond to them⁽¹⁰⁾:

- (2a') دار <dar> *dar* 'house' vs. فار <far> *far* 'mouse' (qualitative distinctive heterography)
 (3a') دار <dar> *dar* 'house' vs. دارو <darw> *daru* 'his house' (quantitative distinctive heterography)
 (4a') كانقرب <kanqrb> *ka-nqarṛab* 'I approach' vs. كان قرب <kan_qrb> *kan qṛab* 'he was closer' (linear distinctive heterography).

Non-distinctive heterography, or in other words *variational heterography*, or simply *variation*,⁽¹¹⁾ is a relation between heterographic words that do not differ in meaning and pronunciation. Such graphic words will be referred to as *graphic variants*. The symbol “~” reads: “is a graphic variant of”. For cases of qualitative variation, the symbol “↔” will be used to mark that two or more graphs are mutually interchangeable in at least some graphic words without this causing a change in their meaning and pronunciation.

Examples (2b), (3b) and (4b) are repeated below as (2b'), (3b') and (4b') as instances of variation, with transcription of phonetic words assumed to correspond to them; in some cases, other graphic variants are added:

- (2b') قال <qal>⁽¹²⁾ ~ كال <kal> ~ گال <ḡal> ~ گَال <ḡal> *gal* 'he said' (qualitative variation)
 (3b') لمغرب <lmgrb> ~ لمغريب <lmgryb> *l-Maḡrib* 'Morocco' (quantitative variation)
 (4b') فالدار <faldar> ~ ف الدار <f_aldar> *fā-d-dar* 'at home, in the house' (linear variation)

Below are given some other typical examples of qualitative (5), quantitative (6) and linear (7) variation:

- (5) داره <darh>⁽¹³⁾ ~ دارو <darw> *daru* 'his house'
 (6) كانوا <kanw> ~ كانوا <kanwa> *kanu* 'they were'
 (7) قالولو <qalwlw> ~ قالو لو <qalw_lw> *qalu lu* 'they told him'

All examples adduced so far illustrate minimal variation: each term of the relation differs from the remaining one (or ones) in a minimal way, i.e. only one distinctive unit is involved. Examples of non-minimal variation, but still representing one of the three types of heterography, are given in (8)-(10):

- (8) أكثر <aktr> ~ أكثر <aktr> *ktar* 'more' (qualitative non-minimal variation)
 (9) لمغرب <lmgrb> ~ المغرب <almgryb> *l-Maḡrib* 'Morocco' (quantitative non-minimal variation)⁽¹⁴⁾

(10) The transcription follows that in Aguadé & Benyahia (2005: 10-12). It is not strictly phonetic and does not use IPA symbols.

(11) The term “variation” is used here in a different way than it was in Michalski (2019: 68-73), where it was reserved for a relation between graphs only.

(12) This form can also represent the pronunciation *qal*, i.e. with [q]. Here, however, the point is that it can represent pronunciation *gal*, with [g].

(13) In this example, this graphy <h> is assumed to represent -u rather than other possible pronunciations.

(14) Some examples, like (9), are repeated – sometimes in a modified form – on purpose in order to better illustrate the particular relations between graphic words.

intended to reflect particular characteristics of MA can be termed self-oriented graphies. A self-oriented graphy may be phonetic, i.e. reflect pronunciation, or morphological, i.e. reflect the morphological structure of a word. The former is illustrated by the use of <t> in تاني <tany> for *tani* ‘second’ and the latter in تدير <tdyr> for *ddir* ‘(that) she does’, with <t> marking the abstract prefix *t-*, realized as *d-* due to assimilation (as opposed to the phonetic graphy دير <dyr>).

WMA makes frequent use of the principle of economy, according to which certain elements of phonetic language may be not represented in script. This can be illustrated by graphies which do not represent vowels and, at the same time, cannot be interpreted as based on any other spelling principle, e.g. مكينش <mkynš> (as opposed to ماكينش <makaynš>) for *ma kayənš* ‘there is not’.

Some graphies are based on what can be termed principle of analogy. This means that a particular graphy is used in graphic word A because the user is aware that it is used in graphic word B which is similar to word A in some respect; however, while this graphy in B is justified by some other spelling principle, this is not the case for A. For instance, the form ابحال <abḥal> for *bḥal* ‘like’ has an unpronounced initial <a>. This graph is used in SA do note a prothetic vowel before consonant clusters, e.g. اكتب <aktb> *uktub* ‘write!’. Thus, in SA its presence is justified by the phonetic principle: it represents a sound. In WMA, the *alif* has no phonetic function, it merely copies the SA graphy because the MA word contains a similar element: the consonant cluster. Another example is the graphy ترايري <tzayry> for *dzayri* ‘Algerian’. Here, the spelling person is aware that *d* is marked in some MA words as <t>, as in تدير <tdyr> for *ddir* ‘(that) she does’ – by virtue of the morphological principle. Consequently, they use <t> by way of analogy, although the analogy between *ddir* and *dzayri* is false: in the latter, *d* is part of the word’s root, not an affix, and the graphy <t> cannot be justified by morphological principle.

2.4 Methodological issues. Corpus

When describing WMA on the Internet, we are faced with some methodological problems which, as a matter of fact, do not substantially differ from those affecting the analysis of WMA in printed literary texts. One fundamental issue is how to determine the intended pronunciation hidden behind a graphic text. Sometimes the basic problem is whether a given graphic expression represents MA or SA – in most cases, context is helpful in this respect. Which particular variety of MA a given text is written in is another question. For instance, does the graph ق <q> in قال <qal> ‘he said’ stand for [q] or [g] and is the word supposed to be read *qal* or *gal*? This can be determined on the grounds of certain other linguistic features manifested in writing but some written forms are indeterminate and should be analyzed adequately (see Michalski 2019: 43-44 and 83-89 for a discussion and examples).

A particular challenge is to differentiate between recurrent, commonly used forms, rare forms and rare, unintentional graphies such as typos, instances of carelessness etc. Without any criterion of what is correct and what is not it is often difficult to decide about it. In general, however, it seems reasonable to assume that the author of a written text makes rational choices based on (i) their knowledge of MA, (ii) the orthography of SA and (iii) the conventions that have so far developed in writing MA. In other words, a graphy may be considered a typo, careless graphy etc. if it is impossible to find a reasonable justification based on these three competences. For instance, it is highly improbable that the following outlandish forms printed in the book *Lə-ṣṣuṣ l-ṣṣayyan* by Ṣzīz ṛ-Rəgrāgi (2008, Rabat) should have been intended by the author:

خودها ذ الجلابة لبسها وديرها ذا الصباط فرجليك (15a)
 <xwdha_d_alğlabh_lbsha_wdyrha_daasbat_frğlyk> (p. 14)

To readers knowing MA, it is obvious what the author intended is: *xud had l-žallaba lbasha u-dir had š-sbat f-rəžlik* ‘take this jellaba, put it on and put on these shoes’ and that the correct graphic form should have the following distribution of spaces:

خوذ هاذ الجلابة لبسها ودير هاذ الصباط فرجليك (15b)
 <xwd_had_alğlabh_lbsha_wdyr_had_alšbat_frğlyk>.

One might expect that printed literary texts, in contradistinction to Internet texts, will show more regular graphy, since they are prepared with more care and are additionally processed by publishing houses. This, however, is usually not the case. The text corpus examined in Michalski (2019), consisting of approximately 1900 printed pages and comprising 32 works by nearly thirty authors, show plenty of variation with a striking amount of creativity, individualism and inconsistency on the part of the authors, resulting in graphies that are strange, not to say extravagant, but sometimes used quite consistently. For instance, in his “novel in the Moroccan language” (*riwaya bə-l-luğa l-məğribiya d-dariža*) titled *ṛ-Rḥil: Dəmša mšāfra* (2012, al-Ribāt: Dār Abī Raqrāq) Murad al-ʿAlami uses (rather inconsistently) atypical graphies such as:

a) marking *i* by means of the sequence <yy>, probably some kind of “etymologizing” spelling, e.g. سويّعات <swyyʿat> *swiʿat* (p. 87) ‘hours (diminutive)’ (cf. **suwayyiʿat*), زوينّة <zwyyʿnh> *zwina* (p. 162) ‘beautiful (fem)’ (cf. **zuwayyina*, with both hypothetical forms following the SA diminutive pattern *CuCaɣyiC*);

b) marking the reduced vowel *ə* as <w>, e.g. إقتول <aqtwl> (p. 21) *yəqtəl* ‘(that) he kills’, يسكون <yskwn> (p. 24) *yəskən* ‘(that) he lives’;

c) marking a perceived assimilation of *n*, i.e. the 1st person present tense prefix, to the stem-initial *r* or *ṛ*, as in اركب <arʿkb> (p. 20) *nərkəb* ‘(that) I ride’, ارجع <arʿʒ> (p. 8) *nəṛʒəf* ‘(that) I return’.

The study Michalski (2019), analyzing MA printed literary texts, is expected to be helpful in the examination of the use of WMA on the Internet, in terms of providing both the terminological apparatus and a comparative perspective. For the needs of the present study, the three following types of Internet texts, representing three different genres, have been selected:

(i) Comments posted by readers (*taʿlīqāt al-zuwwār*) beneath press articles on the “Hespress” online news platform (hespress.com)⁽¹⁸⁾. This text genre is characterized by being created by laymen, ‘ordinary’ Internet users, often having little awareness of how language works; the texts are usually short spontaneous written utterances, sometimes not (proof)read by their authors before being posted and most probably never corrected by anyone. As such, they show a high number of typos and careless graphies.

(ii) Articles in the MA version of Wikipedia (ary.wikipedia.org)⁽¹⁹⁾, the free and open global online encyclopaedia, are a genre on its own: since they are created by a large number of ordinary users who do not necessarily communicate to agree on the issue of graphy, a great

(18) “Hespress” was founded in 2007 and soon became very popular among Moroccans; in 2015 it was the third most-visited site in Morocco (after Google and Facebook) (Roudaby 2015). Apart from MA, the comments are also written in SA, in a combination of both, or in French.

(19) See Sedrati & Ait Ali (2019) for an outline of this project.

variety of forms can be expected. This is, however, counterbalanced by supervision and moderation by the Wikipedia editors.

(iii) Articles in *Gud*,⁽²⁰⁾ an online newspaper publishing texts written partly or fully in MA, represent the genre of journalistic texts. Typically, this genre, being created by professionals and reviewed in an editorial process, is expected to be characterized by a relatively high degree of graphic regularity.

The analysis of the texts representing these genres has consisted in random reading rather than a systematic examination. Whenever necessary, also other sites containing texts written partly or fully in MA have been searched for specific graphic forms with the use of the Google search engine. The queries have been restricted to such sites by selecting those with the Moroccan domain “.ma” or by selecting “Morocco” in the Google “Advanced search” function as well as by verifying that a given text or passage is written in MA.

The analysis has shown that the three genres manifest a number of common graphic features, in particular variation, which they also share with printed literary texts. These similarities are illustrated in Section 3, while Section 4 focuses on the differences. Section 5 is devoted to a comparison between the three genres.

3. Internet texts and printed literary texts: Major shared features

For space limitations, only the most conspicuous graphic features shared by the three genres of Internet texts with printed literary texts in WMA, as described in Michalski (2019) are illustrated here. The following presentation should therefore not be considered exhaustive.

3.1 Qualitative variation.

3.1.1. ا <a> ↔ ة <h> when marking final -a in some nouns, adjectives and participles, e.g.:

- (1) كتوبة <ktwbh> ~ كتبوا <ktwba> *ktuba* ‘books’
- (2) اللولة <allwlh> ~ اللولا <allwla> *l-lūwwla* ‘first (fem)’
- (3) دايرة <dayrh> ~ دايرا <dayra> *dayra* ‘doing (fem)’

3.1.2. ا <a> ↔ ى <á> when marking final -a in some verbal forms, e.g.:

- (4) مشا <mša> ~ مشى <mšá> *mša* ‘he went’
- (5) يبقا <ybqa> ~ يبقى <ybqá> *yībqa* ‘(that) he stays’

3.1.3. ب ↔ پ <p> when marking *p*, e.g.:

- (6) بورطابل <bwrṭabl> ~ پورطابل <pwrṭabl> *poṛṭabl* ‘mobile phone’

(20) *Gud* (in MA, “Straight ahead”), also known in its French transcription as *Goud*, was created in 2011 and “acts as a continuation of «Nichane», a weekly magazine widely read [in the past – M. M.] by Moroccans” (Pennisi 2020a: 85), also owing to a conspicuous presence of MA in its texts.

3.1.4. ك <k> ↔ گ <g> ↔ گ <g> ↔ ج <ǧ> ↔ چ <č> when marking g resulting from the dissimilation of ž in the vicinity of a sibilant,⁽²¹⁾ e.g.:

(7) كالس <gals> ~ گالس <gals> ~ گالس <gals> ~ جالس <ǧals> ~ چالس <čals> ⁽²²⁾ *galas* ‘sitting (masc)’

3.1.5. ك <k> ↔ گ <g> ↔ گ <g> ↔ ق <q> when marking g corresponding to SA q,⁽²³⁾ e.g.:

(8) قال <qal>⁽²⁴⁾ ~ كال <kal> – گال <gal> – گال <gal> *gal* ‘he said’

3.1.6 ك <k> ↔ گ <g> ↔ گ <g> ↔ غ <ǧ> ↔ چ <č>⁽²⁵⁾ when marking g in some foreign words, e.g.:

(9) كوكل <kwkl> ~ گوگل <ǧwǧl> ~ گوگل <ǧwǧl> ~ غوغل <ǧwǧl> ~ چوچل <čwčl> *gugał*⁽²⁶⁾ ‘Google’

3.1.7 ف <f> ↔ ف <v> when marking v, e.g.:

(10) فیديو <fydyw> ~ فیديو <vydyw> *vidyo* ‘video’

3.1.8 In some words, ت <t>, د <d> and ض <ḏ> (phonetic graphy) are interchangeable with ث <ṭ>, ذ <ḏ> and ظ <ẓ> (SA-oriented graphy), respectively, when marking *t*, *d* and *ḏ*, e.g.:

(11) تاني <tany> ~ ثاني <ṭany> *tani* ‘second’

(12) هاد <had> ~ هاذ <haḏ> *had* ‘this’

(13) نضاضر <nḏaḏr> ~ نظاظر <nẓazr> *nḏaḏar* ‘glasses’

3.1.9 The above point is linked to the phenomenon of graphic pseudocorrections (also called hypercorrections), a kind of graphy based on the principle of analogy: the graphs ث <ṭ>, ذ <ḏ> and ظ <ẓ> are used in writing MA words as if following the SA graphy, although the corresponding SA words do not contain the respective sounds *ṭ* ([θ]), *ḏ* ([ð]) and *ẓ/ḏ* ([zʕ]/[ðʕ]). Pseudocorrect graphies, exemplified below on the right hand side are considerably less frequent than phonetic graphies, given on the left, e.g.:

(14) هوت <hwt> ~ حوث <ḥwt> *hut* ‘fish’

(15) غادي <ḡady> ~ غادي <ḡady> *gadi* ‘future verbal particle’

(16) عوض ما <ʕwḏ_ma> ~ عوض ما <ʕwḏ_ma> *ʕwḏ ma* ‘instead of’

(21) On this condition, cf. Heath (2002: 136-138). The use of چ <č> in general is very limited both on the Internet and in printed literary texts (see Aguadé 2005: 247 and 2006: 259 for its use in *Tqarqib n-nab*, a collection of ultrashort texts published by Yusaf Amin l-ʕAlami in 2006).

(22) Brahimi (2024).

(23) Graph غ <ǧ> used to mark this sound is a very infrequent graphy. E.g. حرغت <ḥrǧt> *ḥrǧt* ‘I (illegally) emigrated’ (Ibārūrī 2015).

(24) Cf. footnote 12.

(25) Cf. 3.1.4 above.

(26) «Man takūn Bahīḡa ...» (2020).

3.2 Quantitative variation.

3.2.1 Some vowels –usually those described by linguists as “stable” (Harrell 1962: 10-12) or “long” (Aguadé 2003: 92-93; Moscoso García 2004: 33-36)– can be marked or unmarked in writing. The latter kind of graphy, which is frequently SA-oriented, is referred to as *scriptio defectiva* in Semitic studies. In the former case, they are marked by means of the graphs <a>, <w> and <y>, termed in this function *matres lectionis* and this instance of phonetic graphy is referred to as *scriptio plena*, e.g.:

- (17) <ktāšf> ~ <ktšf> *ktāšaf* ‘he discovered’
- (18) <mṣah> ~ <mṣh> *mṣah* ‘with him’
- (19) <minṭaqḥ> ~ <minṭqḥ> *māṭaqa* ‘area’
- (20) <malyk> ~ <mlk> *malik* ‘king’
- (21) <lmğryb> ~ <lmğrb> *l-məğrib* ‘Morocco’
- (22) <ntwma> ~ <ntma> *ntuma* ‘you (pl)’
- (23) <lwğḥ> ~ <lğḥ> *luğa* ‘language’
- (24) <kwn> ~ <kn> *kun* ‘if’

3.2.2 Sometimes the *scriptio defectiva*, illustrated in the examples below on the right hand side, is used even though it cannot be justified as SA-oriented graphy – because the SA cognate word uses *scriptio plena*, e.g.:

- (25) <kayn> ~ <kyn> *kayən* ‘there is’
- (26) <makaynš> ~ <mkynš> *ma kayənš* ‘there is not’
- (27) <mağadyš> ~ <mğdyš> *ma ġadiš* ‘negated future particle (masc sg)’

Cf. the SA cognates written with *scriptio plena*: <kaʾn> ‘being, existing’, <ma> ‘negation particle’, <ğady> ‘going’.

3.2.3 Gemimates may be marked by means of a single graph (with or without a *šadda* sign) or by a doubled graph, e.g.:

- (28) <nṣs> ~ <nnṣs> *nnṣas* ‘(that) I sleep’
- (29) <ly> ~ <lly> *lli* ‘which’

3.2.4 Gemimates resulting from assimilation can be marked by morphological graphy, i.e. using the graph which represents the abstract, unassimilated unit, e.g. prefix, as is the case with *d-*, corresponding to *t-* on the abstract level, in the following example:

- (30) <dyr> ~ <tdyr> *ddir* ‘(that) she does’

3.2.5 The definite article, pronounced as *l-* or as a consonant resulting from its complete assimilation to the stem-initial consonant *d, ḍ, n, r, s, š, ṣ, t, ṭ, z* or *ẓ*, may be written as <al> (SA-oriented graphy) or <l> (phonetic graphy for *l-* and morphological graphy for other consonants), e.g.:

- (31) <almğryb> ~ <lmğryb> *l-məğrib* ‘Morocco’
- (32) <aldzayr> ~ <ldzayr> *d-dzayər* ‘Algeria’

3.2.6 An initial alif not representing any sound is sometimes written in words beginning with a consonant cluster. If these have SA cognates, such a pseudoprothetic alif can be considered a SA-oriented graphy, e.g.:

- (33) <ally> اللي ~ <lly> للي *lli* ‘which’ (cf. SA <aldy> الذي *alladī* ‘idem’)
 (34) <antwma> انتوما ~ <ntwma> نتوما *ntuma* ‘you (pl)’ (cf. SA <āntm> أنتم ‘you (masc pl)’)

When a MA word has no direct SA cognates, the graphy is based on analogy, e.g.:

- (35) <aʃlaš> اعلاش ~ <ʃlaš> علاش *ʃlaš* ‘why’
 (36) <abḥal> ابحال ~ <bḥal> بحال *bḥal* ‘like’ (preposition)
 (37) <aḥna> احنا ~ <ḥna> هنا *ḥna* ‘we’.

3.2.7 If the preposition *f* ‘in’ is followed by a space, the graphy can be phonetic: ف <f> or SA-oriented: في <fy> (less frequently also فى <fā>, cf. 4.6 below), e.g.:

- (38) <fy_darha> في دارها ~ <f_darha> ف دارها *f-darha* ‘at her home’

3.2.7 Modelled on the *alif al-wiqāya* used in SA graphy to mark some verbal forms, an unpronounced graph ʾ <a> is often added word-finally after و <w> marking the past and present verbal suffixes -u, e.g.:

- (39) <šafwa> شافوا ~ <šafw> شافو *šafu* ‘they saw’
 (40) <yšwfw> يشوفوا ~ <yšwfw> يشوفو *yšufu* ‘(that) they see’

By way of analogy, it is sometimes used after word-final و <w> marking the pronominal suffix -u ‘his, him’, e.g.:

- (41) <ʕndwa> عندوا ~ <ʕndw> عندو *ʕandu* ‘he has’, literally: ‘at/with him’

By the same token, it is used in words simply happening to end in -u, e.g.:

- (42) <walwa> والوا ~ <walw> والو *walu* ‘nothing’
 (43) <hadwa> هادوا ~ <hadw> هادو *hadu* ‘these’

3.3. Linear variation.

3.3.1 Prepositions, usually those composed of one letter, and the conjunction و <w> *w-/u-* ‘and’ can be written with or without a following space, e.g.:

- (44) <byna_w-bynhm> بينا وبينهم ~ <byna_wbynhm> بينا وبينهم *binna w-binhm* ‘between us and them’
 (45) <f_tnḡḥ> ف طنجة ~ <ftnḡḥ> فطنجة *f tanḡa* ‘in Tangier’
 (46) <mn_bʕd> من بعد ~ <mnbʕd> من بعد *mən bəʕd* ‘afterwards’

This also applies to the presentational particles *ha* and *ra*:

- (47) <ha_ḥna> ها حنا ~ <haḥna> هاحنا *ha ḥna* ‘and we...; and here we are’
 (48) <ra_ḥna> را حنا ~ <raḥna> راحنا *ra ḥna* ‘and we...; and here we are’

3.3.2 When a verb or, less frequently, a participle is followed by the preposition *l-* ‘for, to’ with a pronominal suffix attached to it, the whole can be written with or without a space, e.g.:

(49) $\langle ymkn_lyh \rangle \sim \langle ymknlyh \rangle$ *yīmkān lih* ‘he can’, lit. ‘it is possible for him’

(50) $\langle kyban_ly \rangle \sim \langle kybanly \rangle$ *ka-yban li* ‘it seems to me’

(51) $\langle dayra_lyh \rangle \sim \langle dayrallyh \rangle$ *dayra lih* ‘having (fem) done to him’

3.4 Mixed variation.

3.4.1 The negation particle *ma* may be written with or without a following space and with or without an *alif*, e.g.:

(52) $\langle ma_kanš \rangle \sim \langle makanš \rangle \sim \langle mkanš \rangle$ *ma kan-š* ‘he was not’

The same holds true for the present tense preverbs *ka-* and *ta-*, and the future particle *ga-*, e.g.:

(53) $\langle kayqwlw \rangle \sim \langle ka_yqwlw \rangle \sim \langle kyqwlw \rangle$ *ka-ygulu* ‘they say’

(54) $\langle tayqwlw \rangle \sim \langle ta_yqwlw \rangle \sim \langle tyqwlw \rangle$ *ta-ygulu* ‘they say’

(55) $\langle gankwn \rangle \sim \langle ga_nkwn \rangle \sim \langle gnkwn \rangle$ *ga-nkun* ‘I will be’

Forms with a space but no *alif* are infrequent for *ma* (e.g. $\langle m_kanš \rangle$) and hardly used for *ka-*, *ta-* and *ga* (e.g. $\langle k_yqwlw \rangle$, $\langle t_yqwlw \rangle$ and $\langle g_nkwn \rangle$).

3.4.2 The inflexional forms of the verb *qal/gal* ‘to say, to tell’ which end in *-l* and are followed by the preposition *l-* ‘to, for’ with a pronominal suffix attached to it can be written with one ل <l> or two لـ <ll> and with or without a space between the verb and the prepositional phrase, e.g.:

(56) $\langle qalyha \rangle \sim \langle qal_lyha \rangle \sim \langle qallyha \rangle$ *gal-liha* (or *qal-liha*) ‘he told her’ (lit. ‘he said to her’)

(57) $\langle qwlyna \rangle \sim \langle qwl_lyna \rangle \sim \langle qwllyna \rangle$ *gul-lina* (or *qul-lina*) ‘tell us’ (lit. ‘say to us’)

4. Internet texts and printed literary texts: Major differences

Generally, it can be said that in comparison with printed literary texts, Internet texts written in MA contain more self-oriented graphy (first and foremost phonetic graphy, including *scriptio plena*) and less SA-oriented forms (see Section 5 for details). Pseudocorrections and extravagant graphies appear to be less frequent. Obviously, such a comparison should be understood as one based on the relative frequency of use of particular solutions rather than their being exclusive to one category or the other.

Arguably, the differences result from the fact that the corpus of Internet texts in WMA, significantly larger than that of printed literary texts, is being processed by large masses of users on a daily basis. Therefore, it favours more universal solutions and, consequently, some kind of spontaneous standardization.

Some most conspicuous examples of such spontaneously (semi-)standardized features of the Internet texts, differentiating it, in general terms, from the graphy of the printed literary texts, are discussed in what follows.

4.1 Predominance of the phonetic graphy of the pronominal suffix *-u* ‘him, his’ و <w> over the SA-oriented graphy ه <h>.

4.2 Only marginal use of ه <h> to mark the word-final *a* in nouns, adjectives and participles, such as كبيره <kbyrh> *kbira* ‘big (fem)’ (cf. 3.1.1.).

4.3 Use of the bare alif ا <a> word-initially, in various functions, rather than إ <ā>, آ <ā> or إ <a>, which are all quite frequent in printed literary texts.

4.4 Absence of vocalization signs, with the occasional use of ؤ <ʾ> (the *šadda*).

4.5 No use of the pseudoprothetic *alif* in words not beginning with a consonant cluster, e.g. ابلي <abllly> *bəlli* ‘that’ (conjunction) or انتسناك <antsnak> *nətsənnak* ‘(that) I wait for you’ (cf. Michalski 2019: 104).

Not in every respect, however, do Internet texts look more regular in comparison with printed literary texts. The following two points illustrate this.

4.6 In addition to ي <y>, its undotted counterpart ى <ā> is quite often used to mark the word-final *i* in various parts of speech, e.g.:

- (1) مغربي <mğrby> ~ مغربى <mğrbā> *mğarbi* ‘Moroccan’
- (2) يمشي <ymsy> ~ يمشى <ymsā> *yīmši* ‘(that) he goes’

In the Internet texts, ى <ā> is also quite often used with no phonetic function in one of the two SA-oriented graphies of the preposition *f* ‘in’, viz. فى <fā> (the other one being the more frequent في <fy>). In literary texts, it was identified in one source only (see Michalski 2019: 181-183 for more details).

4.7 The digraph دج <dğ> is used to mark [dʒ] in borrowings from English, e.g. الدجاز <aldğaz> *l-ğaz* [ldʒæz], especially in proper names transcribed into MA, e.g. دجون <dğwn> *ğon* [dʒon] ‘John’. Such graphies have, however, more frequent variants with ج <ğ>: الجاز <alğaz>, جون <ğwn>. In the corpus of printed literary texts, no relevant occurrences have been identified.

The above list is by no means exhaustive. The types of differences that can be identified between the orthography used in printed literary texts and on the Internet depend largely on the corpora selected as the basis of investigation. The corpus referred to in Section 2.4 appears sufficiently large to provide reliable comparative data. By contrast, the Internet corpus of WMA texts is so vast that a more thorough examination would require the application of statistical tools. Nevertheless, even on the basis of random sampling and less quantifiable observations, it is evident that the spontaneous standardization occurring in this medium, as compared to WMA in printed literary texts, is striking. This tendency, however, does not extend equally across all areas of the Internet or across all text genres. These issues will be discussed in greater detail in the following section.

5. Types of variation and text genres

In this Section, the three genres are characterized with respect to their most conspicuous graphic features, with a focus on variation. A comparison of their graphies will make it possible to assess which one of them shows which degree of regularity.

5.1. Readers' comments

Readers' comments beneath the texts published in the "Hespress" online news platform represent spontaneous, usually very short texts. Created by people who, as a rule, do not live by their pen and do not tend to feel a need to check and correct what they have written, such pieces contain a relatively high number of typing errors, i.e. forms not intended by the author, as well as forms difficult to justify linguistically, sometimes even internally inconsistent. For instance, the word *tbərgig* 'inquisitiveness-cum-gossiping', which contains two sounds *g*, can be written consistently as, for instance, <tbrgyg>⁽²⁷⁾, but it is not difficult to find internally inconsistent form such as, e.g. <tbrgyk>⁽²⁸⁾, with two different graphs, گ <g> and ك <k>, marking one sound. Likewise, the word *gugəl* 'Google' has such aleatory variants as <gwk>⁽²⁹⁾ <kwğl>⁽³⁰⁾, <ğwğl>⁽³¹⁾ as well as other unjustified forms. The word *nḍaḍər* 'glasses', which contains two sounds *ḍ*, can be written, apart from other forms, as <nḍazr>⁽³²⁾ or <nḍaḍr>⁽³³⁾.

An example of linguistically unjustified linear features is the use a space after ي <y> representing the inflectional prefix *y-/yī-*, illustrated below:

- (1) <ymkn_baqy_ky_šwf_bšynyh>⁽³⁴⁾ (instead of <kyšwf> كيشوف) *yīmkən baqi ka-yšuf b-šinih* 'maybe he is still seeing with his eyes'
- (2) <ky_dyr_mbga>⁽³⁵⁾ (instead of <kydyr> كيدير) *ka-ydir ma bga* '(he) does what he wants'

The latter example also contains a relevant quantitative feature: the negation particle *ma* written as م <m>, rather than ما <ma>. This can be interpreted as an instance of the principle of economy at work: minimizing efforts at writing, or typing, is a paramount factor determining graphy in this genre. Other examples:

- (3) <mkyṅš> *ma kayənš* 'there is no' (~ <makayṅš> ماكايṅش ~ <ma_kayṅš> ما كايṅش)
- (4) <mğdš> *ma ġadiš* 'negated future particle' (~ <mağadyš> ماغاديš ~ <ma_ğadyš> ما باغيš)
- (5) <mbğyš> *ma bağiš* 'do(es) not want' (~ <mabagyš> ماباغيš ~ <ma_bağyš> ما باغيš)
- (6) <mkyṅnš> *ma kaynin-š* 'there are no' (~ <makayṅnš> ماكايṅنيš ~ <ma_kayṅnš> ما كايṅنيš)

(27) Eg. Yass (2017).

(28) E.g. Daxəl Suq Rašu (2018).

(29) E.g. caprice (2022).

(30) E.g. Ġarīr (2016).

(31) E.g. Rašīd (2012).

(32) E.g. Salīm (2017).

(33) Simo Asfi (2016).

(34) Muṣṭafā (2019)

(35) Anas (2024).

Consider also the two examples of the minimalistic graphy ل <l> for *lli* ‘which’:

- (7) <ana_l_m_fhamtš>⁽³⁶⁾
ana lli ma fhəmt-š
 ‘what I didn’t understand’

- (8) <hdšy_l_šta_all’h>⁽³⁷⁾
had š-ši lli šta llaḥ
 ‘This is what God has given’

Economy seems also to be the reason for a rather limited use of the special graphs پ <p> for *p*, ف <v> for *v*, and گ <g> or ك <g> for *g*. They are not in directly accessible in typical Arabic keyboard configurations and most users do not like to waste their time looking for them. In this case, the principle of economy would thus have the upper hand over the phonetic graphy.

The two next instances show how the phonetic graphy is given preference over the morphological one:

- (9) <mmbšd_tlatyn_šam>⁽³⁸⁾
mim bəšd tlatin šam (cf. *min bəšd* in careful pronunciation)
 ‘after thirty years’

- (10) <ma_šty_walw>⁽³⁹⁾
ma šatti walu (cf. *ma šaṭti walu* in careful pronunciation)
 ‘you saw nothing’

The idea of ‘write as you speak’, put into practice in these two graphic utterances by reflecting the assimilation of *n* to *m* and of *f* to *t*, is one that permeates the genre under discussion.

5.2 Wikipedia in MA

If one considers that Wikipedia, “the free encyclopædia”, is created by amateurs and ordinary users and that a single article may be written or edited by more than one author or editor, a multitude of disparate graphies amounting to a graphic chaos can be expected. This is, however, not the case because, in Wikipedia, a text’s author is not the sole person responsible for its quality: Thanks to the positive controlling role of the Moroccan Wikipedia editors, called *imḡarən* (plural from *amḡar*, which is an Amazigh word for ‘chief, leader’), its articles show a strikingly high degree of regularity and low number of extravagant forms. In addition, the Moroccan Wikipedia has published in its *Kūnnaš lə-qwāšad* ‘Book of rules’⁽⁴⁰⁾ a number of explicit rules concerning language (*lə-ḥkam də l-luḡa*), both grammar and graphy, of obligatory (*ilzami*) or recommendatory (*tužihi*) character. Additionally, following a general practice (no relevant rule has been explicitly announced, to the best of my knowledge) supervised by the editors, an article should be written in one variety of MA, in order to avoid dialectal and stylistic

(36) Fouad (2021).

(37) Muḥammad Nīhū Šəkrād (2018).

(38) Nūmīdī Amāzīgī Dzayrī (2018).

(39) Chwarla (2018).

(40) *Kūnnaš lə-qwāšad* (no date).

incoherence⁽⁴¹⁾. From this, it follows that the users are not required to use a specific variety of MA when initiating an article; what matters is consistency within the scope of a single article.

The most relevant aspects of orthography listed in the *Kūnnaš lā-qwāṣad* are summarized in what follows.

Its Rules 1 to 3 concern, in general, the use of a space (called *xwa*⁽⁴²⁾) after prepositions and conjunctions. They stipulate that a space must be used after:

- a) genitival prepositions دِيَال <dyaḷ> *dyaḷ*, د <d> *dā*, تَاع <taṣ> *taṣ*, مَتَاع <mtaṣ> *mtaṣ* and نَتَاع <ntaṣ> *ntaṣ*⁽⁴³⁾ (rule 1).
- b) conjunctions listed as وَ <w> and وَ <w̄> for *w-/u-* ‘and’⁽⁴⁴⁾ and وَلَا <wla> for *wūlla* ‘or’ (rule 3)⁽⁴⁵⁾.

Also the remaining prepositions (called *l-kālmāt dā ḡ-ḡarṣ*) should be separated from the following word by a space, including those consisting of one graph, e.g. ب حُرُوف <b_ḥrwf> *bā-ḥrurf* ‘with letters’, ل حَبَابِكَ <l_ḥbabk> *lā-ḥbabāk* ‘for your loved ones’ (Rule 2). This rule is, however, recommendatory.

For *-a* marking the feminine nominal ending, it is recommended (Rule 4) that nouns which “come from an Arabic word” (*ḡayin mən kālma ḡarḡbiyya*) – i.e. have a SA equivalent – written with final ة <ḥ> should also be written in this way (including in a genitival construction, when final *-a* is replaced in pronunciation with *-t*). Borrowings from languages other than Arabic (*s-smīyyat* “*l-gawriya*”) can be written either with ة <ḥ> or ا <a>. The former is recommended for words in which the final *-a* is pronounced short, exemplified with مِيطَة <myṭḥ> *miṭa* ‘half’ (in sports) (from French *mi-temps*), كَاسِيطَة <kasyṭḥ> *kaṣiṭa* ‘cassette’, and بَلَاصَة <blaṣḥ> *blaṣa* ‘place’ (from Spanish *plaza* ‘place’), while the latter is preferable if *a* is pronounced longer, as in مَالِيزِيَا <malyzya> *malizya* ‘Malaysia’, إِثْنُولُوجِيَا <aṭnwlwḡyā> *itnoloḡya* ‘ethnology’ and جِيُولُوجِيَا <ḡywlwḡyā> *ḡyoloḡya* ‘geology’.⁽⁴⁶⁾

The spelling of the definite article (*t-taḡrif*) (Rule 5) is regulated as follows: If a word begins with a non-assimilating consonant, the article *l-* should be written as ل <l>, e.g. لِابَاب <lbab> *l-bab* ‘the door’. If the first consonant is an assimilating one, the article can be marked either as ال <al>, e.g. الدَار <aldar> *d-dar* ‘the house’, or as a *ṣadda* (gemination sign) on the first letter, e.g. دَّار <ḍar>. The rule is recommendatory, in contrast to one stipulating that only the option with ال <al> is allowed if a word functions as title of an article (the *ṣadda*-equivalent can be used for redirection).

(41) I owe this latter piece of information to Mounir Afifi (Wikimedia Morocco User Group), personal communication (26 April 2025).

(42) All rules are formulated in MA, showing an early stage of the development of its linguistic terminology.

(43) Called *ḡrūf r-rāḡt u-l-wūṣl* there, more literally: ‘particles of joining and linking’. Note that they are dialectally differentiated.

(44) No explanation of the difference between these two forms is provided and their use is illustrated in the same context: لَمَغْرِب و دَزَايِر <lmḡryb_w_d:zayr> and لَمَغْرِب وَ دَزَايِر <lmḡryb_w̄_d:zayr> *l-Maḡrib u-d-Dzayar* ‘Morocco and Algeria’.

(45) It is also signalled, rather superfluously, that combinations with suffixed pronouns, e.g. دِيَالِي <dyaḷi> *dyaḷi* ‘min’, literally: ‘of mine’, are written with no space.

(46) Perhaps the issue should not (only) be considered in terms of the vowel’s length but a particular word’s morphology being or not being (perceived as) compatible with the structure of typical MA words.

Finally, the use of the hamza graph, marking glottal stop [ʔ], word-finally is described in Rule 6. If a MA word does not contain this sound, the hamza is not written, even though it is in its SA cognate with this sound, e.g. ما <ma> *ma* ‘water’ (cf. SA ماء <mao> *māʔ*). Its use is, however, obligatory in those MA words in which it is always pronounced, as in إنشاء <aṇṣao> *inšaʔ* ‘foundation, creation’.

Although no rules on marking the typically Moroccan sounds, [g], [p] or [v] are specified in the Moroccan Wikipedia, instructions (sometimes incomplete) are given on its help-page on how to introduce certain special graphs: گ <ġ>, گ <ġ>, پ <p> and the *šadda*.⁽⁴⁷⁾ A search of its pages for some particular forms with the use of the “Search for pages containing...” function (“*Qallab fla sfahi fihum...*”) shows that these special graphs are preferred to phonetically ambiguous ordinary Arabic letters, viz. ك <k> marking *g* or ب marking *p*. For instance, the name *Mbappé* is written مبابي <mbapy>, with پ <p>, on all seven pages mentioning this French football player, while the form مبابي <mbaby>, with ب , is not used at all (it redirects to مبابي <mbapy>). The form راپ <rap> for *rap* ‘rap (music)’ occurs twice as frequently as راب <rab> and the word *gaʿ* ‘all’ is written كاع <kaʿ>, with the phonetically ambiguous ك <k> on 18 pages, while the graphy گاع <ġaʿ> occurs on 347 pages (for گاع <ġaʿ>, 5 pages are found).

An interesting feature of the Moroccan Wikipedia graphic system is the quite consistent use of the digraph دج <dġ> for [dʒ] in words of English origin (see Section 4.7).

There is a visible tendency to mark long, or stable, vowels with the *scriptio plena*, although there is no regulation on it. For instance, the full name of Wikipedia, *l-māwsuʿa l-ḥurra* ‘the Free Encyclopaedia’, is written in its logo as لموسوعة لحررا <lmwswʿh_l-ḥwra>, not حرا <ḥra> (or حرة <ḥrh>). A search of Moroccan Wikipedia’s pages shows that there is marked preference for the *scriptio plena* in the case of certain lexemes, e.g. (the values in brackets indicate the number of pages with a given graphy):

- | | |
|---|---|
| (1) <i>luġa</i> ‘language’ | : لوغة <lwġh> (635) vs. لغة <lġh> (183) |
| (2) <i>l-Maġrib</i> ‘Morocco’ | : لمغريب <lmġryb> (6,383) vs. مغرب <lmġrb> (60) |
| (3) <i>d-dariža</i> ‘the colloquial (language)’ | : الداريجة <aldaryġh> (558) vs. الدارجة <aldarġh> (234) |

In some other lexemes, however, the long vowels tend not to be marked with graphs, e.g.

- | | |
|--|---|
| (4) <i>l-ʿilm</i> ‘knowledge, science’ | : لعيلم <lʿylm> (11) vs. لعلم <lʿlm> (31) |
| (5) <i>l-mašriq</i> ‘East’ | : لمشرق <lmšriq> (2) vs. لمشرق <lmšrq> (11) |
| (6) <i>l-ḥurr</i> ‘free (definite)’ | : لحرور <lḥwr> (3) vs. لحر <lḥr> (13) |

The unpronounced SA-oriented *alif al-wiqāya* in certain verbal forms ending in *-u* is quite consistently avoided. For instance, *kanu* ‘they were’ is written without it as كانوا <kanw> on more than 2,000 pages, while the graphy كانوا <kanwa> is used on 30 only. Similarly, *ykunu* ‘(that) they are’ is written without it, as يكونو <ykwnw>, on 325 pages, whereas the form يكونوا <ykwnwa> has been found only on one page. The use of the pseudo-prothetic *alif* is an exception rather than a rule. Both these choices to favour the self-oriented phonetic graphy and consistence with which they are put into practice are yet another indication that the Moroccan Wikipedia takes the issue of a consistent graphy seriously.

(47) *Mṣawna* (no date).

5.3 Online newspaper *Gud*

Newspapers can be assumed to publish texts created by professionals and persons guided by professionals. It can also be expected that their editorial staff include someone responsible for taking care of the graphic homogeneity of the texts. An investigation of the content of the *Gud*, however, does not validate these expectations: its texts show a high level of irregularity. The general impression is that its creators give the authors leeway as to what kind of graphy they want to use. To illustrate this: when looking at one of the articles,⁽⁴⁸⁾ the reader can see three different ways of marking the sound *g* at a time: in the newspaper's logo, گ <g> is used in its name *Gud*, written گود <gwd>. One of the thematic sections listed at the top of the page is called *Tbargig*, roughly: 'society gossip', with both its *g*-sounds represented as ك <k>: تبركك <tbrkyk>. In the title of the article, this sound is marked as گ <g> in the words: *gaʕ* 'all', *ħarṛaga* 'illegal emigrants', and *ħarḡu* 'emigrated illegally'.

Similarly, there is little consistency as far as quantitative features are concerned: Forms with *alif al-wiqāya* seem not to be used as often as those without it, but nevertheless they are quite frequent. Likewise, the pseudo-prothetic *alif* is not an uncommon sight. There are lexemes for which the *scriptio plena* is quite common, e.g. موهيم <mwhym> for *muhimm* (although the SA-oriented form مهم <mhm> is more usual), but in certain words it is not as frequent as, for instance, in the Moroccan Wikipedia. For example, for *d-dariža* 'colloquial (language)' written الداريجة <aldarġh> there are nearly 300 hits, while الداريجة <aldaryġh> has only three; *l-Maġrib* 'Morocco' is written as المغرب <almġrb> or, less often, as لمغرب <lmġrb>, with the alternatives المغرب <almġryb> and لمغرب <lmġryb> being marginal or non-existent. Perhaps these particular words, and some others, are written with SA graphy in mind.

Linear variation is quite strong. It can be exemplified by the graphy of the preposition *f*- 'in'. In the title of one article,⁽⁴⁹⁾ the phrase *f-Balžika* 'in Belgium' is written with a space: ف بلجيكا <f_blžyka>, whereas no space is used in فالاستفادة <falastfadħ> *fə-l-istifada* 'in the profit'. In the opening sentence of the text, however, the latter word is written with a space: ف الاستفادة <f_alastfadħ>. In other articles, there are yet other graphies of *f-Balžika*: with no space: فيلجيكا <fblžyka>⁽⁵⁰⁾ or in the SA-oriented form في بلجيكا <fy_blžyka>⁽⁵¹⁾ (cf. SA في <fy> for *fī* 'in').

In her analysis of the role of the *Gud* in the development of WMA, Pennisi (2020a: 93-94) observed that "[i]t is also thanks to editorial experiences such as «Goud», that *dāriġah* undertook (and undertakes) the process of informal codification (see conventionalization or «standardization from below») through its extensive written use." From our above observations, however, one gets the impression that the newspaper does not seem to have made substantial achievements as far as standardization of graphy is concerned. The present observable usage makes it rather doubtful that any conscious efforts towards elaborating a consistent graphy are being undertaken in its editorial environment.

6. Conclusion

The description provided by the present work may seem incomplete as it lacks some statistical underpinning. Its intention, however, has not been an exhaustive quantitative exploration but rather signalling the most important phenomena and indicating major paths of

(48) «S-suluṭāt l-maġribiya...» (2024).

(49) «157 mġarbi...» (2024).

(50) E.g. in «Ilyās š-Šāḡar...» (2024).

(51) E.g. in š-Šūfi (2023).

investigation. Several issues could not be treated here for space considerations, for instance the graphy of emphatic consonants, the initial *hamza* and sounds corresponding in MA borrowings from French to nasal vowels as well as the use of non-final forms of graphs before a space, such as ف فاس rather than فاس ف.

In comparison with printed sources, viz. books, Internet texts show a range of features that make them a better exponent of the processes currently taking place in the universe of WMA: the rapid growth of their volume means that, on one hand, more and more phenomena are reflected in writing and, on the other hand, an increasing number of people have access to it and are exposed to it. Consequently, it also has a greater flexibility in adjusting itself to the users' needs. As such, its role in shaping any future norm could be considerably more important than that of printed (literary) texts – especially if its development is not left run wild but overseen at least in some respects. In other words, it is where spontaneity meets control that the greatest potential of WMA on the Internet for shaping the future of its graphy lies.

Symbols and abbreviations

MA	– Moroccan Arabic
SA	– Standard Arabic
WMA	– Written Moroccan Arabic
~ (between graphic words)	– is a graphic variant of
↔ (between graphs)	– is mutually interchangeable in (at least some) qualitative variants with

Bibliography

Primary Sources

(all sites accessed 31.03.2025)

- «157 mgərbi ...» (2024) = «157 mgərbi həslu f-Bəlžika bə-sbab lə-htiyāl ʕla l-hukuma fə-l-istifada mən s-sakan l-ižtimāʕi», *Gud*, 02.04.2024, <https://www.goud.ma/157--مغربي-حصلو-ف-بلجيكا-بسباب-الاحتيا-ل-ع-871619>
- ANAS (2024): [Hūm (!) dāʕiman yubarrirūn...], reader's comment, *Hespress*, 28.12.2024, <https://www.hespress.com/1489192-أسعار-حلو-ي-بوناني-تسجل-الارتفاع-والم-871619>
- BRAHIMI, Youssef (2024): «Umarāʔ al-ğahīm wa-l-sabʕ wa-l-xaṭāyā», *Critique*, 29.02.2024, <https://critique.ma/أمراء-الجحيم-السبع-و-الخطايا-871619>
- CAPRICE (2022), [Akbar ʕamaliyyat našb...], reader's comment, *Hespress*, 27.04.2022, <https://www.hespress.com/980099-جمهورية-إفريقيا-الوسطى-تعتمد-البتكو-871619>
- CHWARLA (2018): [L-ixwan ana bgit r-rədd...], reader's comment, *Hespress*, 13.10.2018, <https://www.hespress.com/455882-مؤشر-دولي-لحالة-الجوع-يرمي-بالمغرب-إلى-871619>
- DAXƏL SUQ RAŠU (2018), [Waš baqi ʕayšin...], reader's comment, *Hespress*, 10.02.2018, <https://www.hespress.com/409725-أصغر-رئيسة-جماعة-بالمغرب-نفي-الخيانة-871619>
- «Dṛəg n-nšāra»...» (2023) = «Dṛəg n-nšāra» yuʕarqil tanaqqul al-muwāṭinīn bi-Sīdī Binnūr», *al-Usbūʕ*, 10.08.2023, <https://www.alousboue.ma/94357/>
- FATHĪ, Nihād (2018a): «L-ʕar la-ma xəlli w Saʕd lə-Mžərrəd yittəšhər bə-ffaylu maši bə-ğnah», *Gud*, 28.08.2018, <https://www.goud.ma/377984-العار-لأما-خليو-سعد-المجرد-يتشهر-بفعالي-871619>
- FATHĪ, Nihād (2018b): «L-islamiyun kəffsuha ʕla l-mužtamaʕ...», *Gud*, 26.07.2018, <https://www.goud.ma/371197-الإسلاميين-كفسوها-على-المجتمع-و-ما-ع-871619>
- FOUAD (2021): [Ana lli ma fhəmt-š...], reader's comment, *Hespress*, 07.01.2021, <https://www.hespress.com/761997-الكونغرس-يصادق-على-فوز-بايدن-بالانتخاب-871619>

- ĠARĪR (2016): [Xtaṛəf lina ši ḥaḥa ūḫra...], reader's comment, *Hespress*, <https://www.hespress.com/292555-بلخياط-ينتقد-هيمنة-غول-وقيسبو.html>
- ĠARBI, Asma (2016): «L-qərqubi u-ma ydir. Bi-ṣ-ṣuwar...», *Gud*, <https://www.goud.ma/255655-القرقوبي-ومايدير-بالصور-اعتقال-شاب-شر/>
- «Honkong» (no date), *Wikipedia*, <https://ary.wikipedia.org/wiki/هونكونغ>
- IBĀRŪRĪ, Sihām (2015): «Təhlil nəfsi-iḥtimaḥi l-haduk lli mṣəddḥinna...», *Gud*, 29.01.2015, <https://www.goud.ma/119559-تحليل-نفسى-اجتماعى-لهادوك-اللى-مصدقين-/>
- «Ilyās š-Šāḥər...» (2024) = «Ilyās š-Šāḥər tṭəḥkəm ḥlih bə-l-ḥəbs f-Bəlḥika», *Gud*, 23.02.2024, <https://www.goud.ma/864268-الياس-الشاعر-تحكم-عليه-بالحبس-فبلجيكا-/>
- Kūnnaš lə-qwāḥəd (no date), https://ary.wikipedia.org/wiki/ويكيبيديا:كناش_لقواعد
- Mṣawna (no date): https://ary.wikipedia.org/wiki/معاونة:صفحة_لمعاونة
- «Man takūn Bahīḡa ...» (2020) = «Man takūn Bahīḡa Ḥāfiḥ allatī ḥtafā Gügl bi-ḍikrā mawlidihā al-112?», *al-Axbār*, 04.08.2020, <https://www.alakhbar.press.ma/-من-تكون-بهيجة-حافظ-التي-111014-احتفى-جوجل-بذكر.html>
- MUḤAMMAD NĪHŪ ŠƏKRĀD (2018): [Ahl al-Məḡrib mən ahli...], reader's comment, *Hespress*, 22.04.2018, <https://www.hespress.com/422782-مشارك-ملكية-في-مهب-الريح.html>
- MUṢṬAFĀ (2019): [Yimkən baqi ka-yṣuf...], reader's comment, *Hespress*, 1.04.2019, <https://www.hespress.com/488012-الرئاسة-الجزائرية-بوتفليقة-سيستقيل-ق.html>
- MZĪYYƏN, ũUmər (2016): «Fḍiḥa f-Ḍar lə-Brihi : Ilyas l-ũUmari ṛḥəf küllu fazəg bə-l-ḥərq», *Gud*, 27.07.2016, <https://www.goud.ma/232438-فضيحة-فدار-لبريهي-الياس-العماري-ارجع-ك.html>
- NŪMĪDĪ AMĀZĪĠĪ DZAYRĪ (2018): [Anā ka-Nūmīdī...], reader's comment, *Hespress*, 04.08.2018, <https://www.hespress.com/443040-نشطاء-أمازيغ-يطلقون-حملة-تيفيناغ-عل.html>
- RAŠĪD (2012): [Idā kuntum taṣtabirūn al-luḡa...], reader's comment, *Hespress*, 23.05.2012, <https://www.hespress.com/233-88399-جمعية-تطالب-البرلمان-بالتراجع-عن-نع.html>
- SALĪM (2017): [Lā arā ayyat niswa...], reader's comment, *Hespress*, 13.07.2017, <https://www.hespress.com/370867-بنمسعود-حكومة-بنكيران-مانعت-تفعيل-مسا.html>
- SIMO ASFI (2016): [Šəft gaḥ l-baṣar bḡa...], reader's comment, *Hespress*, 21.03.2016, <https://www.hespress.com/278990-النظارايتيون-يخوضون-اضرابا-وطنيا-ويرف.html>
- «S-Si Bənkiran...» (2012) = «S-Si Bənkiran: hada wūlla tḥərhiš», *Gud*, 11.02.2012, <https://www.goud.ma/96145-2-السي-بنكيران-هدا-ولا-تبرهش.html>
- Š-ŠŪFĪ, Karīm (2023): «Təḍḥiya kbira: 2 mḡarba f-Balḥika tlaḥu fə-myah konḗele baš...», *Gud*, 09.11.2023, <https://www.goud.ma/842604-تضحية-كبيرة-2-مغاربة-في-بلجيكا-تلاحف.html>
- «S-suluṭāt l-məḡribiya ...» (2024) = «S-suluṭāt l-məḡribiya qəblat l-ḥəwda dyalhūm...», *Gud*, 29.03.2024, <https://www.goud.ma/870883-السلطات-المغربية-قبلات-العودة-ديالهم.html>
- YASS (2017): [Anā attafiq maṣak fī kull mā qult...], reader's comment, *Hespress*, 14.12.2017, <https://www.hespress.com/399224-أخصائي-مخدر-النميمة-يرفع-منسوب-الكر.html>

Secondary Sources

- AGUADÉ, Jordi (2003): «Estudio descriptivo y comparativo de los fonemas del árabe dialectal marroquí», *Estudios de dialectología norteafricana y andalusí*, n° 7, pp. 59-109.
- AGUADÉ, Jordi (2005): «Darle al pico: un 'bestiario' de Youssouf Amine Elalamy en árabe marroquí», *Estudios de dialectología norteafricana y andalusí*, n° 9, pp. 245-265.
- AGUADÉ, Jordi (2006): «Writing dialect in Morocco», *Estudios de dialectología norteafricana y andalusí*, n° 10, pp. 253-274.
- AGUADÉ, Jordi (2012): «Monarquía, dialecto e insolencia en Marruecos: el caso *Nichane*», *De los manuscritos medievales a internet: la presencia del árabe vernáculo en las fuentes escritas*, Zaragoza: Universidad de Zaragoza, pp. 441-464.

- AGUADÉ, Jordi (2013): « Des romans diglossiques: le cas de Youssef Fadel », *Évolution des pratiques et représentations langagières dans le Maroc du XXI^e siècle*, vol. 1, Paris: L'Harmattan, pp. 207-220.
- AGUADÉ, Jordi & BENYAHIA, Laila (2005): *Diccionario árabe marroquí: Árabe marroquí-español, español-árabe marroquí*, Cádiz: Quorum Editores.
- BENÍTEZ FERNÁNDEZ, Montserrat (2003): «Transcripción al árabe marroquí de mensajes de teléfono móvil», *Estudios de dialectología norteafricana y andalusí*, n° 7, pp. 153-163.
- BENÍTEZ FERNÁNDEZ, Montserrat (2010): *La política lingüística contemporánea de Marruecos: de la arabización a la aceptación del multilingüismo*, Zaragoza: Instituto de Estudios Islámicos y del Oriente Próximo.
- BENÍTEZ FERNÁNDEZ, Montserrat (2012a): «*TelQuel*: una fuente contemporánea para el estudio del árabe marroquí», *De los manuscritos medievales a internet: la presencia del árabe vernáculo en las fuentes escritas*, Zaragoza: Universidad de Zaragoza, pp. 403-417.
- BENITEZ FERNANDEZ, Montserrat (2012b): «*Al-ʔAmal*: Otro intento fallido de escribir en *dārīza* marroquí», *Dynamiques langagières en Arabophonies: Variations, contacts, migrations et créations artistiques. Hommage offert à Dominique Caubet par ses élèves et collègues*. Zaragoza: Universidad de Zaragoza, pp. 379-391.
- BERJAOUI, Naser (2001): «Aspects of the Moroccan Arabic orthography with preliminary insights from the Moroccan computer-mediated communication», *Chat-Kommunikation: Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation: Perspektiven auf ein interdisziplinäres Forschungsfeld*, Stuttgart: Ibidem, pp. 431-465.
- CAUBET, Dominique (2004): « L'intrusion des téléphones portables et des 'SMS' dans l'arabe marocain en 2002-2003 », *Parlers jeunes ici et là-bas: Pratiques et représentations*, Paris: L'Harmattan, pp. 247-270.
- CAUBET, Dominique (2005): « Génération *darija*! », *Estudios de dialectología norteafricana y andalusí*, n° 9, pp. 233-243.
- CAUBET, Dominique (2012): « Apparition massive de la *darija* à l'écrit à partir de 2008-2009: sur le papier et sur la toile: Quelle graphie? Quelles régularités? », *De los manuscritos medievales a internet: la presencia del árabe vernáculo en las fuentes escritas*, Zaragoza: Universidad de Zaragoza, pp. 377-402.
- CAUBET, Dominique (2013): « Maroc 2011 – Messagerie instantanée sur l'internet marocain: *facebook*, *darija* et parlers jeunes », *Évolution des pratiques et représentations langagières dans le Maroc du XXI^e siècle*, vol. 1. Paris: L'Harmattan, pp. 63-87.
- CAUBET, Dominique (2017): «Morocco: An informal passage to literacy in *dārīja* (Moroccan Arabic)», *The politics of written language in the Arab world: Writing change*, Leiden-Boston: Brill, pp. 116-141. https://doi.org/10.1163/9789004346178_007
- ELINSON, Alexander E. (2013): «*Dārīja* and changing writing practices in Morocco», *International Journal of Middle East Studies*, n° 45, 4, pp. 715-730. <https://doi.org/10.1017/S0020743813000871>
- FERRANDO, Ignacio (2012): «Apuntes sobre el uso del dialecto en la narrativa marroquí moderna», *De los manuscritos medievales a internet: la presencia del árabe vernáculo en las fuentes escritas*, Zaragoza: Universidad de Zaragoza, pp. 349-358.
- GÜNTHER, Hartmut (1988): *Schriftliche Sprache: Strukturen geschriebener Wörter und ihre Verarbeitung beim Lesen*. Tübingen: Max Niemeyer Verlag. <https://doi.org/10.1515/9783110935851>
- HARRELL, Richard S. (1962): *A short reference grammar of Moroccan Arabic*. Washington, D. C.: Georgetown University Press.
- HEATH, Jeffrey (2002): *Jewish and Muslim dialects of Moroccan Arabic*. London-New York: RoutledgeCurzon.

- HOOGLAND, Jan (2013a): «Towards a standardized orthography of Moroccan Arabic based on best practices and common ground among a selection of authors», *Árabe marroquí: de la oralidad a la enseñanza*, Cuenca: Ediciones de la Universidad de Castilla-La Mancha, pp. 59-76.
- HOOGLAND, Jan (2013b): «L'arabe marocain, langue écrite», *Évolution des pratiques et représentations langagières dans le Maroc du XXI^e siècle*, vol. 1. Paris: L'Harmattan, pp. 175-188.
- HOOGLAND, Jan (2018): «Darija in the Moroccan press: The case of the magazine *Nichane*», *Sociolinguistic Studies*, n° 12, 2, pp. 273-293. <https://doi.org/10.1558/sols.35567>
- LANGONE, Angela Daiana (2003): «*Hbār blādna*. Une expérience journalistique en arabe dialectal marocain». *Estudios de dialectología norteafricana y andalusí*, n° 7, pp. 143-151.
- MGARFAWI, Xalīl & MABRUR, ʿAbd al-Wāḥid & ŠUKAYRI, ʿAbd Allāh (2017): *Qāmūs al-dāriġa al-maġribiyya* [Dictionary of Moroccan Darija], al-Dār al-Bayḍā: Markaz Tanmiyat al-Dāriġa Zakūra.
- MICHALSKI, Marcin (2016): «Spelling Moroccan Arabic in Arabic script: The case of literary texts». *Arabic varieties – Far and wide: Proceedings of the 11th International Conference of AIDA – Bucharest, 2015*, București: Editura Universității din București, pp. 385-394.
- MICHALSKI, Marcin (2019): *Written Moroccan Arabic: A study of qualitative variational heterography*, Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- AL-MIDLĀWĪ AL-MNABBĤĪ, Muḥammad (2019): *Al-ʿArabiyya al-dāriġa: Imlāʾiyya wa-naḥw – al-aṣwāt, al-ṣarf, al-tarkīb, al-muṣğam* [Colloquial Arabic: Spelling and grammar – sounds, morphology, syntax, lexis], al-Dār al-Bayḍā: Markaz Tanmiyat al-Dāriġa Zakūra.
- MILLER, Catherine (2012): «Observations concernant la présence de l'arabe marocain dans la presse marocaine arabophone des années 2009-2010», *De los manuscritos medievales a internet: la presencia del árabe vernáculo en las fuentes escritas*, Zaragoza: Universidad de Zaragoza, pp. 419-440.
- MILLER, Catherine (2015): «Des passeurs individuels au mouvement linguistique: itinéraires de quelques traducteurs au Maroc», *Le social par le langage – La parole au quotidien*, Tunis: IRMC – Paris: Karthala, pp. 203-232.
- MILLER, Catherine (2017): «Contemporary *dārija* writings in Morocco: Ideology and practices», *The politics of written language in the Arab world: Writing change*, Leiden – Boston: Brill, pp. 90-115. https://doi.org/10.1163/9789004346178_006
- MOSCOSO GARCÍA, Francisco (2004): *Esbozo gramatical del árabe marroquí*. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.
- MOSCOSO GARCÍA, Francisco (2009): «Comunidad lingüística marroquí en los foros y chats: Expresión escrita, ¿norma o anarquía?», *Al-Andalus Magreb*, n° 16, pp. 209-226.
- MOUSTAOU SRHIR, Adil (2016): *Sociolinguistics of Moroccan Arabic: New topics*. Frankfurt am Main: Peter Lang. <https://doi.org/10.3726/978-3-653-06484-1>
- PENNISI, Rosa (2020a): «Expressions of resistance: “Goud” and stylistic variation in Moroccan digital newspapers», *La rivista di Arablit*, n° X, 20, pp. 79-98.
- PENNISI, Rosa (2020b): «Written *dārija*: “māšī māsḡūl taktāb-ha bi-ḥurūf al-luġa al-ʿarabiyya!” It is not logical to write it with the Arabic letters! Media reception of the Zakoura Dictionary project», *Annali di Ca' Foscari. Serie orientale*, n° 56, pp. 129-153. <https://doi.org/10.30687/AnnOr/2385-3042/2020/01/005>
- PENNISI, Rosa (2023): «From Mixed Arabic to Educated Written *dāriġa*: Diglossic variation in Moroccan written production», *Revista Española de Lingüística*, n° 53, 2, pp. 293-314. <https://doi.org/10.31810/rse.53.2.11>

ROUDABY, Youssef (2015): « Comment *Hespress* s'est hissé au sommet », *TelQuel*, 12.03.2015 (https://telquel.ma/2015/03/12/comment-hespress-hisse-sommet_1438110, accessed 31.03.2025).

SEDRATI, Anass & AIT ALI, Abderrahman (2019): «Moroccan *Darija* in online creation communities: Example of Wikipedia», *Al-Andalus Magreb*, n° 26. (<https://revistas.uca.es/index.php/aam/article/view/6794/6744>).
<https://doi.org/10.25267/AAM.2019.i26.11>