ARABIC VARIETIES: FAR AND WIDE

Proceedings of the
11th International Conference of AIDA
Bucharest, 2015

editors
GEORGE GRIGORE
GABRIEL BIȚUNĂ

Arabic Varieties: Far and Wide

*Proceedings of the 11th International
Conference of AIDA – Bucharest, 2015*

Arabic Varieties: Far and Wide

Proceedings of the 11[th]International Conference of AIDA – Bucharest, 2015

EDITORS

George Grigore
Gabriel Bițună

CONTENTS

SPELLING VARIANTS IN WRITTEN EGYPTIAN ARABIC,
A STUDY ON LITERARY TEXTS

LUCIA AVALLONE

The University of Bergamo
Foreign Languages, Literatures and Communication Studies

Abstract: Egyptian Arabic is the spoken vernacular for the most populated Arab country. It is also a form of written communication in informal genres, for instance chats, blogs, emails etc., and in more traditional genres such as novels, short stories, poetry, and theatre. Although the current systems of online communication have seen it emerge as a relevant phenomenon in recent years, vernacular writing is not a novelty in the modern age. From the last decades of the nineteenth century onwards, Egyptian literature has offered several cases of works written in vernacular. This paper presents some results of a study carried out, in a diachronic perspective, on literary texts belonging to different genres and discourse modes, in search of words written according to different spellings to identify variants, to compare them, and to evaluate their common and diverse elements. Indeed, by comparing the practice of writing Egyptian Arabic which has not been codified as a literary means, some noteworthy features emerge not only from the lexical and morphological choices of authors but also from the graphic representation of this language variety, allowing the description of a framework of variants which could be considered as a basic corpus in a possible operation of normalizing the vernacular orthography.

Keywords: *Vernacular, diachronic perspective, variants, graphic representation, normalization.*

Introduction

A sharp distinction between the two main varieties of Arabic made on the basis of the communication channel used has been a landmark in the linguistic norm transmission and an essential postulate for the main authorities of the Arabic language, such as academies and much of the Arab intellectual milieu[1]. However it does not represent the real language use, which is far more complex. In fact, such a strict division of roles, which provides for the adoption of the Arabic *fuṣḥā* exclusively in writing, by limiting the use of the *'āmmiyya* (or *dāriǧa*) to the spoken, is not in force now and it has not been in the past.

Studying the effects generated in writing the vernacular can make a contribution to describe the composite framework of linguistic uses in the Arabic-speaking community. Among the results of writing the vernacular, there is a not uniform process of adapting an oral variety to the Arabic script. The paper here presented aims to introduce some outcomes of a research in the field of the Egyptian vernacular graphic representation, with a particular interest for the adaptation phenomena[2].

---

[1] Šawqī Ḍayf, elected president of the Academy of the Arabic Language in Cairo in 1996, in his paper entitled *al-'Āmmiyya fuṣḥā muḥarrafa* explains the factors which make the *fuṣḥā* a proper means for the Arab civilization and the *'āmmiyya* a sub-product merely apt to the ordinary oral communication (Ḍayf 2000: 34). Among the intellectuals who give a strong support for *fuṣḥā* there are Naǧīb Maḥfūẓ and Ṭāhā Ḥusayn. They consider it the language of writing and fulfilling the Arab society speculative purposes: "I adopted *fusha* [when I started writing] because it was the [accepted] language of writing. The question [of *fusha* and *'ammiyya*] has become problematic only in relatively recent times. Many people consider it a serious problem, and it may well be so in the theatre or cinema. But in the novel and short story, it is much less serious and time alone will settle the question" (Maḥfūẓ 1977: 61, in El-Enany 1993: 193); "I am, and shall remain, unalterably opposed to those who regard the colloquial as a suitable instrument for mutual understanding and a method for realizing the various goal of our intellectual life because I simply cannot tolerate any squandering of the heritage, however slight, that classical Arabic has preserved for us. The colloquial lacks the qualities to make it worthy of the name of a language. I look upon it as a dialect that has become corrupted in many respects" (Hussein 1998: 89).

[2] As Egyptian vernacular, like other Arabic language varieties, lacks a standard orthography, the transition from the oral realization to the written one implies a spontaneous spelling which could be transformed into a conventionalized orthography

The written texts selected for this study belong to modern Egyptian literature and cover a period of time ranging from the late nineteenth to the twenty-first century. The diachronic perspective answers to the purpose of detecting if there have been specific trends in the practice of transliterating the oral variety into the written one, through a spelling adjustment based on norms established by custom, if some uses have been typical of some periods, and if a codification of the vernacular writing is now conceivable.

In the past writing dialect caused debates (Daniëls 2004) among linguists, writers and, more broadly, intellectuals[3], but they did not change the basic distinction between a high literary production and a "popular literature"[4] or, at most, a literary production placed in an undefined area. Actually, texts completely written in Egyptian Arabic, which have had a significant impact in Arabic literature, have been treated as products outside the norm which have not paved a way nor changed the tendency of the majority of authors, from the late nineteenth century throughout the twentieth century, to write in standard Arabic. But the 2000s do show a stronger presence of Egyptian literary works in the vernacular (Rosenbaum 2010, Avallone 2011), which implies the scholars of literature need a greater understanding of these trends, while for linguists it has become more and more interesting to observe the graphic representation modes, an operation that cannot be done without querying the genesis and development of this aspect.

1. Methodology

The research intends to examine texts chosen to cover most of the period and taken from what we might call the "canon" of the Egyptian vernacular literature, that means works entirely written in the vernacular that have had success in the literary environment, and texts which, though belonging to the official literary canon, present dialogues in the vernacular. As here only the results pertinent to the first stage of the study are presented, the corpus taken into consideration in this paper is formed by works completely or mostly written in Egyptian vernacular.

I have assumed that the vernacular has been gradually defined also as a written and a literary language, and that the variety of Cairo, already prestigious as a spoken language, has established itself as a written model (Rosenbaum 2004: 283). Although the need has not so far been felt to set norms of the vernacular writing or, at least, it has not been translated into reality, who writes in *'āmmiyya* refers to a number of conventions related to phonological, morphological, and lexical aspects, developed over time and accepted by authors and readers. The hypothesis is that the various graphic transcription features could be leaning towards uniformity.

The work has been structured in six phases apt to focus on some aspects –quantifiable and comparable linguistic elements– in order to understand the trend of the single texts, and to compare between the different authors: 1. choosing texts belonging to the "canon" of the Egyptian vernacular literature; 2. choosing a priori a sample for each text; 3. reading the samples and identifying high frequency grammatical items which have variants; 4. recording the identified elements to quantify their occurrence within the single samples; 5. analysing uniformity or variation in the elements spelling within the single samples; 6. comparing the results arising from the different samples.

(Eskander et al. 2013) and which involves a series of adaptation phenomena at lexical, phonological and morphological levels.

[3] In the context of the language choice between standard and vernacular Arabic, perceived as linked to the modern literary genres development, the debates which take place on the pages of periodicals *al-Muqtaṭaf* in the years 1881 and 1882, and *al-Hilāl* in 1949 are significant. In the first case the voices debating the subject identify in the standard the variety most accredited to represent modernity in writing. In the second one, which involves leading writers, the role of the vernacular is emphasized as a means of renewal of literature and as a variety more suited to the dramaturgical representation.

[4] The terms 'popular' and 'folk' are broad and inclusive of texts belonging to different genres which share the feature of being addressed to a wide audience that recognize and adopt them in their own identity and heritage. I mention here an excerpt from the *Encyclopædia Britannica* definition of "popular literature": "Popular literature includes those writings intended for the masses and those that find favour with large audiences. It can be distinguished from artistic literature in that it is designed primarily to entertain. Popular literature, unlike high literature, generally does not seek a high degree of formal beauty or subtlety and is not intended to endure".

2. Corpus

Sample 1 [S1] *Ṣuḥuf Abū Naḍḍāra* by Yaʿqūb Ṣannūʿ. Five articles (sketches), twenty-two pages (pp. 1-12; 26-27; 30-32; 34-35; 38-40). Satirical journal, 1878.

Sample 2 [S2] *al-Ustāḏ* by ʿAbd Allāh an-Nadīm. Four articles (sketches), twenty-one pages (pp. 46-48; 65-70; 132-140; 147-149). Satirical journal, 1892.

Sample 3 [S3] *es-Sayyid we-mrāto fi Maṣr* by Bayram at-Tūnisī. Eight chapters, thirty-four pages (pp. 7-40). Satirical dialogue, 1925.

Sample 4 [S4] *Muḏakkirāt ṭālib biʿṭa* by Luwīs ʿAwaḍ. One chapter, twenty pages (pp. 33-52). Autobiography, 1942, pub. 1965.

Sample 5 [S5] *Qanṭara allāḏī kafara* by Muṣṭafā Mušarrafa. One chapter, thirty-two pages (pp. 5-36). Novel, 40s, pub. 1966.

Sample 6 [S6] *Riḥla fī n-Nīl* by Ṣabrī ʿUṯmān. Two chapters, twenty-eight pages (pp. 88-115). Humorous long tale, 1965.

Sample 7 [S7] *Laban al-ʿuṣfūr* by Yūsuf al-Qaʿīd. Three chapters, thirty-five pages (pp. 8-42). Novel, 1994.

Sample 8 [S8] *Marāʿī l-qatl* by Fatḥī Imbābī. Four chapters, twenty-six pages (pp. 7-32). Novel, 1994.

Sample 9 [S9] *Tāksī* by Ḫālid al-Ḫamīsī. Nine chapters, thirty-eight pages (pp. 13-50). Fictional dialogues, 2006.

Sample 10 [S10] *ʿAyza atgawwez* by Ġāda ʿAbd al-ʿĀl. Eight chapters, thirty-eight pages (pp. 5-42). Long tale based on a blog, 2008.

3. Results

The items identified to carry out an analysis of the single texts and a comparison among them are morphological (prepositions, demonstratives, future tense markers, preverbal particle *bi-*, personal pronouns, constructions consisting of verb followed by *li-*+pronominal suffix) and lexical (active participle of the verb meaning 'to want', adverb of time meaning 'now', conjunction meaning 'also', negation particle meaning 'not'). In addition, other elements have emerged as worth noting, but only pertaining to some texts, for instance lexical items containing glottal stops derived from the phoneme /q/ and words ending with *alif maqṣūra* or their variants with a final *alif*.

    The data collected point out that, generally speaking, authors' attitudes in respect to either the adoption of vernacular spelling norms previously elaborated by other writers or the coining of new modalities to be coherently applied in their texts are very changeable. Some samples display a trend towards unambiguous choices and others a plurality of forms for the same word.

3.1. [S1] and [S2]

The oldest texts analysed, [S1] and [S2], date back to the late nineteenth century. They show a different practice of rendering Egyptian vernacular in Arabic script. Ṣannūʿ seems to follow norms quite near to the standard script, in most cases but not always, without deviating from them, while an-Nadīm's sketches are characterized by a plurality of forms for some words.

    On the whole, in [S1] and [S2] we note that the prepositions *fi*, *min*, *ʿala*, and *bi-* (this latest also preverbal particle) occur in their canonical standard forms with only some exceptions for *ʿala*: [S1] $<في>_{185}$ [5], $<من>_{173}$, $<على>_{120}$ vs. $<ع>_1$, and $<بـ>_{126}$; [S2]] $<في>_{74}$, $<من>_{68}$, $<على>_{72}$ [6] vs. $<علـ>_5$ vs. $<ع>_1$, and $<بـ>_{80}$. There are several examples of words occurring with final long vowel /ā/ or desinence /ah/: demonstratives *da* and *keda* occur in both their widespread forms in [S2], $<دا>_{22}$ and $<ده>_{11}$, $<كده>_5$ and $<كدا>_2$, while in [S1] $<ده>_{75}$ is prevalent on $<دا>_3$, and $<كدا>_{23}$ effectively prevailing on $<كده>_1$, a

---

[5] The subscript indicates the occurrence of each item in the sample.
[6] The preposition على also occurs in the adverbial expression $<على شان>_1$, usually written as a compound ($<علشان>$).

rarer form <كيده>$_1$ occurring as well. As to [S2], <براً>$_2$ and <بره>$_3$ are both registered. In [S1] *ḥatta* always occurs with *alif maqṣūra* <حتى>$_{13}$; the verb *baqa*[7] is realized as <بقي>$_{17}$ and <بقا>$_5$; third person masculine pronouns have the forms <هو>$_9$ and <هم>$_2$, but also the plural variant <هما>$_4$. Further to the spelling duplicity of final phonemes, examples in [S2] are: <عنده>$_2$ vs. <عندو>$_1$ and <هو>$_8$ vs. <هوه>$_8$. [S2] attests the adverb meaning 'also' in two spellings: <برده>$_4$, including an example with suffix –*hum*, and <برضه>$_1$. In Ṣannūʿ's texts only <برضه>$_1$ occurs. Besides, the two samples show some demonstrative compounds: [S1] <دا أنا>$_3$, <دانتم>$_1$; [S2] <دنا>$_1$, <دانا>$_1$, and <دنتي>$_1$. In [S1] the negative particle occurs with the spelling <مش>$_7$, while in [S2] the long form <موش>$_{14}$ is registered.

What typifies an-Nadīm's third sketch is the large amount of words containing a glottal stop which in standard Arabic writing corresponds to the grapheme <ق> and which is mostly written here using the *hamza*. So we find many verb 'to say' items represented through the *hamza* with the three kinds of support or without any, instead of the letter *qāf* which occurs only in <تقول>$_4$ and <قلت>$_1$, and, as to the verb 'to be/to become', in <يبقى>$_3$. With the exception of the word *maqlūba*$_1$, there are several nouns written with the *hamza*, such as <ارش>$_1$, <عألّه>$_1$, <لؤمه>$_4$, <منّص>$_1$, <وائف>$_1$ or verbs such as <لئينا>$_1$, <تنرا>$_1$, and <تنعد>$_2$, all this in contrast to the other two sketches of [S2] which show a more usual spelling through the grapheme <ق>. This feature is completely absent in the Ṣannūʿ's texts I have analysed.

A final remark is to be made about verbs followed by preposition *li-*+pronominal suffix which can be written as separated elements or as compounds. [S1] shows one case of such compounds, <احكيلك>$_1$, and some examples of the disjointed construction: <قال لي>$_6$, <احكيها لكم>$_1$, <احكي لي/لنا>$_2$, <قل لي>$_1$, <يقول له>$_1$. In [S2] both the possibilities are applied more extensively, in fact we find the verb 'to say' forms <تقول لي>$_4$ and <قلت لي>$_1$, as well as <بآلك>$_1$, <آتلو>$_1$, <أألكم>$_1$, <أألك>$_1$, and other verb items such as <خلينالك>$_1$, <يخليلك>$_1$, <يخليهالك>$_1$, <يجيلهم>$_1$, <يجيلك>$_1$, <يجولو>$_1$, and <اجيب لك>$_1$ and <يجيبو لنا>$_1$.

3.2. [S3]

The dialogue published in the 20s by Bayram at-Tūnisī, *es-Sayyid we-mrāto* [S3], displays a sharp prevalence of some forms in respect to others, testifying a tendency towards definite author choices. Though some items are registered in two forms, for instance <هو>$_{14}$ vs. <هوه>$_7$, <انت>$_5$ vs. <انته>$_3$ and <هم>$_4$ vs. <همه>$_1$, other are mostly found in their canonical form: prepositions <على>$_{40}$ vs. <عـ>$_{11}$, <من>$_{49}$ vs. <مـ>$_1$; only one spelling is attested for <في>$_{85}$ and <بـ>$_{116}$; demonstrative <ده>$_{20}$ is favoured in respect to <دا>$_8$ and only <كده>$_{35}$ occurs, not <كدا>. The future tense particle is found as <حا>$_{13}$ and <حـ>$_{10}$, but not as <هـ>; the negative particle occurs as <مش>$_{35}$, the adverb *dilwaqt(i)* as <دلوقت>$_{19}$ and the active participle <عايز>$_{11}$ is prevalent on <عاوز>$_1$. The tendency towards uniformity is clear also in the representation of verbs followed by preposition *li-* and pronominal suffixes: 21 disjointed constructions vs. 6 compound constructions (<بقولك>$_3$, <ماقولكيش>$_1$, <قايللي>$_1$, <فكرولكم>$_1$).

3.3. [S4] and [S5]

As to the 40s, the two texts analysed, [S4] by ʿAwaḍ and [S5] by Mušarrafa, show some common features in morphological items. The prepositions *min*, *fi*, and *ʿala* occur in both canonical and short forms: [S4] <من>$_{37}$ vs. <مـ>$_1$, <في>$_{68}$ vs. <فـ>$_{24}$, <على>$_{19}$ vs.<عـ>$_{15}$; [S5] <من>$_{74}$ vs. <مـ>$_{33}$, <في>$_{108}$ vs. <فـ>$_{73}$, <على>$_{101}$ vs. <عـ>$_{45}$. The demonstrative masculine pronoun/adjective *da* occurs in two different spellings: [S4] <دا>$_{16}$ vs. <ده>$_3$; [S5] <ده>$_{18}$ vs. <دا>$_6$. With respect to differences, in [S4] the preposition and preverbal particle *bi-/b-*[8] occurs either as a prefix or as an isolated form, <بـ>$_{35}$ vs. <ب>$_{13}$ and in [S5] only the spelling <بـ>$_{194}$ is attested. Besides, in [S5] the demonstrative item, in short

---

[7] The grapheme <ق>, realized according to the Cairene pronunciation as [ʾ], is here indicated as *q* to recall the correspondent standard letter.

[8] [b-] before 1st person singular.

or long form (<د> and <دا>), occurs also with personal suffixes: <دنا>$_7$, <دانتي>$_1$, <دانها>$_1$, and <دانه>$_1$; the demonstrative *keda* in [S5] is registered exclusively in the variant <كده>$_{23}$, while in [S4] it is documented in both variants <كده>$_4$ and <كدا>$_2$; the future tense particle occurs either as a prefix or an isolated form, with two possible phonological realizations, /ḥa/ and /ha/, and different spellings: [S4] <ح>$_3$, <حـ>$_1$, <حا>$_1$, <هـ>$_3$, and <ها>$_2$, [S5] <حـ>$_{12}$, and <حا>$_1$; the negative particle *miš/muš* occurs in one form in [S4], as <مش>$_{32}$, but in [S5] it is recorded as <مش>$_{19}$ and also as <موش>$_6$ and <مووش>$_2$; the active participle of the verb meaning 'to want' is documented as spelled in two forms in [S4], <عاوز>$_7$, <عايز>$_2$ (feminine <عاوزه>$_1$ and <عايزاني>$_1$, plural <عاوزين>$_3$), and in one form in [S5] <عاوز>$_9$ (feminine <عاوزه>$_2$ and plural <عاوزين>$_1$); the lexical item *dilwaqt(i)* occurs as <دلوقت>$_8$ and <دولوقت>$_1$ in [S4] and <دلوقتي>$_2$ and <دلوقت>$_3$ in [S5].

3.4 [S6]

The sample attests the author's leanings towards a certain uniformity, although not fully reached. This text is particularly interesting for its long introduction written in Egyptian vernacular concerning Ṣabrī's theory of the new Arabic language, that is *ʿāmmiyya*, adoption. It is worthy of note that this introduction regards both the link between language choice and ideology, and the reformist thought.

No shortened forms of prepositions *ʿala*, *fi* and *min* occur and not even isolated forms of preposition and preverbal particle *bi-* or future tense particle *ḥa-*. Demonstratives are registered only as <ده>$_{36}$ and <كده>$_{18}$, the first occurring also as <دانت>$_2$ and <دانه>$_1$ in compound forms. One occurrence of pronoun <ذلك>$_1$ is registered.

As to choices related to lexical items, some significant features are found: the negative particle is only <مش>$_{21}$; the active particle of the verb meaning 'to want' occurs twice as <عاوزين>$_2$; the adverb meaning 'now' is mostly spelled as <دلوقت>$_{13}$ with one exception <دلوقتى>$_1$; the preposition meaning 'for', 'for the sake of' occurs in one form, <علشان>$_{13}$; some words such as *emta*, *ḥatta*, and *baqa* are written with a final *alif* not *maqṣūra* as happens in other texts. The preposition *ʿala* occurs in both variants <على>$_9$ and <علا>$_{83}$. The adverb meaning 'inside' is spelled as <جوه>$_1$ and <جوا>$_1$. Finally it is interesting to note that the verb *qāl* never forms a compound with a following preposition *li-*+pronominal suffix. Two occurrences of a comparable compound are registered, <اشترولي>$_1$ and <يجيبو لي>$_1$, though the second item cannot be considered sure due to lack of uniform typographical features: <yigībūlī> or <yigībū lī>?

3.5. [S7]

The data collected from *Laban al-ʿuṣfūr* point out as a significant feature the coexistence of prepositions with their short homologous (<على>$_{40}$ and <ع>$_{19}$, <في>$_{80}$ and <ف>$_{28}$, <من>$_{69}$ and <م>$_{11}$), and the occurrence of the future tense particle as <حا>$_{26}$ and <حـ>$_4$, in addition to a major occurrence of <دا>$_{25}$ vs. <ده>$_{10}$. The active participle meaning 'wanting' occurs as <عاوز>$_8$, <عايز>$_4$, <عاوزين>$_2$, <عايزة>$_5$, <عاوزة>$_5$ and <عايزاهم/ها>$_3$ (feminine participle+pronominal suffix). Other items occur in two variants: <عشان>$_{24}$ vs. <علشان>$_2$; <هوه>$_{27}$ vs. <هو>$_{13}$, <هيه>$_8$ vs. <هي>$_5$, and <همه>$_8$ vs. <هم>$_1$. A univocal representation of *bi-* (<بـ>$_{132}$), *keda* (<كده>$_{20}$), and *muš/miš* (<مش>$_{20}$) is relevant. As to the verb 'to say' followed by preposition *li-*+personal suffixes, it occurs in both constructions, separated or linked: <لهم قال>$_1$, <قالت لي>$_1$, <قال لى>$_1$, <قلت للي>$_5$, <قلت له/لهم>$_6$, <يقول له/لي/ليها>$_8$, <لك/له/لها/لهم أقول>$_2$. <يقوللي>$_1$, <تقولي>$_1$, <قايلي>$_1$, <قاللي>$_6$, <قالي>$_1$; <قالو له>$_1$.

3.6. [S8]

Marāʿī l-qatl is characterized by the occurrence of few preposition and particle variants: <على>$_{131}$ vs. <ع>$_{11}$; <من>$_{192}$ vs. <م>$_5$. The preposition *bi-* is represented only as <بـ>$_{152}$ and the future tense particle as <حـ>$_{42}$. Some vernacular items are spelled according to the official language writing practice, for

instance personal pronouns and words ending with *alif maqṣūra* such as *ḥatta* <حتى>$_{18}$ and *emta* <امتى>$_2$ (≡standard <متى>). The negative particle occurs only as <مش>$_9$. It is worth noting that the sample also contains grammatical items typical of standard Arabic, for instance demonstratives as *hāḏā* and *ḏalika* vs. the vernacular <ده>$_7$, <دا>$_6$ and <كده>$_2$, and relatives as *allātī* and *allāḏī* (instead of the vernacular <اللي>). The active participle meaning 'wanting' occurs as <عايز>$_8$, <عايزة>$_3$, and <عايزين>$_1$. The occurrence of verbs followed by *li-* and personal suffixes is not substantial but testifies the adoption of a disjointed construction.

3.7. [S9] and [10]

The two texts written in the 2000s, [S9] by al-Ḫamīsī and [S10] by ʻAbd al-ʻĀl, display shared modalities of spelling: the preposition and preverbal particle *bi-* occurs only in prefix form, with cases of vowel lengthening in combination with pronominal suffixes ([S9] <بيهم>$_2$ and <بيه>$_1$; [S10] <بيها>$_3$ and <بيه>$_6$). The prepositions *ʻala*, *fi* and *min* occur in long form in both texts ([S9] <على>$_{55}$, <في>$_{93}$, <من>$_{47}$; [S10] <على>$_{64}$, <علا>$_2$, <في>$_{134}$, <من>$_{71}$), likewise in [S10] the occurrence of short forms <ع>$_{25}$ and <م>$_9$ is significant (<ف>$_1$ is also attested); in [S10] the lengthening of vowel /i/ is attested with preposition *li-*+pronominal suffix, for instance <ليهم>$_2$. The demonstrative pronoun/adjective *da* is registered in both [S9] and [S10] in only one form ([S9] <ده>$_{25}$; [S10] <ده>$_{61}$); in [S10] two compound forms of <دا> occur, <دانا>$_2$ and <دانه>$_2$. The demonstrative *keda* occurs in the form <كده> in both texts ([S9] <كده>$_9$, [S10] <كده>$_{61}$) and <كدة>$_2$ in [S9]; no occurrences of <كدا> are documented. In [S9] the personal pronoun variants <همه>$_1$ and <هما>$_2$ occur. A compound form of emphatic particle *ma-* joined to the personal pronoun *ana* occurs in [S10] <مانا>$_1$. The spellings of the future tense particle show different choices on the part of the authors, as al-Ḫamīsī uses two forms, the isolated one <ح>$_9$ and the prefixed one <حـ>$_{20}$, while in ʻAbd al-ʻĀl's sample we find all the future verbs prefixed with the particle <هاـ>$_{40}$ or <هـ>$_6$. The active participle meaning 'wanting' is documented as <عاوز>$_1$ and <عايز>$_2$ (also plural <عايزين>$_2$ and feminine <عايزة>$_1$) in [S9], while in [S10] the only form registered, though with genre and number variants, is <عايز>$_8$ (plural <عايزين>$_1$ and feminine <عايزة>$_{15}$). The lexical item *dilwaqt(i)* occurs as <دلوقت> and <دلوقتي> in both samples with a prevalence of the second form ([S9] <دلوقتي>$_4$ vs. <دلوقت>$_1$; [S10] <دلوقتي>$_9$ vs. <دلوقت>$_1$); the adverb meaning 'also', 'too', occurs with the spelling <برضه>$_2$ in [S9] and as <برضه>$_{12}$ and <برده>$_1$ in [S10]; the preposition meaning 'for', 'for the sake of' occurs in two forms, *ʻašān* and *ʻalašān*, in both texts, but with different prevalence ([S9] <علشان>$_9$ vs. <عشان>$_5$; [S10] <عشان>$_{25}$ vs. <علشان>$_1$). Two diverse ways of writing the verb 'to say' followed by the preposition *li-* introducing a pronominal indirect object are showed: the verb can be written separated from the indirect object, for instance <قال لي>, like in standard Arabic, or in a compound form, <قاللي>. Examples are: [S9] <قال لى>$_6$, <قال لك>$_1$, <قالوا لي>$_1$, قال له>$_3$, <قلت له>$_4$, <لنا قالو>$_2$, and <أقولك>$_2$, <تقوللي>$_2$, <تقوله>$_2$, <يقولك>$_1$, <قاللي>$_3$, <قاله>$_1$; [S10] <قلت له>$_3$, <هاتقولولي>$_1$, <تقوليله>$_1$ etc. An analogous treatment is found with the verb 'to tell' linked to a following pronominal indirect object introduced by the preposition *li-*, <احكيلك>$_1$, in [S9].

4. Discussion

The data collected make a major aspect emerge: the variety of choices between a canonical (≡standard) graphic representation and a freer (≈vernacular) one, either for prepositions, or demonstratives, or verbs followed by preposition *li-*+pronominal suffix.

As to prepositions, in [S1] and [S2] we note canonical standard forms, with only some exceptions for *ʻala* which occurs also as a prefix <عـ>$_1$ in [S1] and <علـ>$_5$ in [S2]. In [S3] prepositions are mostly found in their canonical form, with the variants <عـ>$_{11}$ (vs. <على>$_{40}$) and <م>$_1$ (vs. <من>$_{49}$). In [S4] the preposition and preverbal particle *bi-/b-* occurs either as a prefix <بـ>$_{35}$ or in isolated form <ب>$_{13}$; the prepositions *fi*, *ʻala*, and *min* are documented in both long and short forms (<في>$_{68}$ vs. <ف>$_{24}$; <على>$_{19}$ vs. <ع>$_{15}$; <من>$_{37}$ vs. <م>$_1$). Also in [S5] both long and short forms of prepositions *fi*, *ʻala*, and *min* are registered, with a sharp prevalence of the long one (<في>$_{108}$ vs. <ف>$_{73}$; <على>$_{101}$

vs.<ع>$_{45}$; <من>$_{74}$ vs. <م>$_{33}$). As said previously, analogous instances are found also in [S7] and [S8], while in [S9] only the long forms are attested. It is worth mentioning that in [S10] some short forms of prepositions occur but only for *'ala* we notice a relevant occurrence, though not a prevalence, <على>$_{64}$ vs. <ع>$_{25}$.

As to demonstratives, in [S1] *da* with final *hā'* <ده>$_{75}$ is effectively prevalent on <دا>$_{3}$ with final *alif*. On the contrary, the final *alif* <كدا>$_{23}$ prevails on the final *hā'* <كده>$_{1}$. In [S2] the distribution comparison is not so dramatic, maybe because the relevant items registered are fewer; at any rate it points out the two different practices: <دا>$_{22}$ vs. <ده>$_{11}$ and <كده>$_{5}$ vs. <كدا>$_{2}$. Demonstrative compounds are found in both samples but transliterated according to different modalities: <دا أنا>$_{3}$ and <دانتم>$_{1}$ in [S1] and <دنا>$_{1}$, <دانا>$_{1}$, and <دنتي>$_{1}$ in [S2]. In [S3] both forms for *da* are attested, with a prevalence of <ده>$_{20}$ vs. <دا>$_{8}$, while *keda* occurs only in one form, with final *hā'* <كده>$_{35}$. Also [S4] and [S5] testify the occurrence of the two different spellings for *da*: <دا>$_{16}$ vs. <ده>$_{3}$ in [S4] and <ده>$_{18}$ vs. <دا>$_{6}$ in [S5]. *Keda* is spelled only as <كده>$_{23}$ in [S5], while it occurs in both variants <كده>$_{4}$ and <كدا>$_{2}$ in [S4]. [S6] shows a univocal representation of both demonstratives <ده>$_{36}$ and <كده>$_{18}$, a usage found also in [S9] and [S10], and partially in [S7] and [S8] where only one spelling of <كده> occurs but *da* is registered in both forms: <دا> vs. <ده>.

As regards verbs followed by preposition *li-*+pronominal suffix, the items found in the samples point out the fact that their graphic representation is rather free. Verbs such as 'to say', 'to tell', and 'to give' can be written as compounds with preposition *li-* followed by pronominal suffix, a choice pertaining to an-Nadim whose texts show disjointed forms <اجيب لك>$_{1}$, <قلت لي>$_{1}$, <تقول لـي>$_{4}$, and <يخليهاللك>$_{1}$, <يجيلهم>$_{1}$, <يجيلك>$_{1}$, <يجولو>$_{1}$, <أألك>$_{1}$, <أألكم>$_{1}$, <آنلو>$_{1}$, <بآلك>$_{1}$, but also <يجيبو لنا>$_{1}$, <احكيلك>$_{1}$, <خلينالك>$_{1}$, and <يخليلك>$_{1}$. In Ṣannū''s texts instead only one case of such compounds occurs, and in Ṣabrī's sample two occurrences have been registered, <اشترولي>$_{1}$ and <يجيبو لي>$_{1}$. In al-Qa'īd's sample, as to the verb *qāl*, there are some instances: the construction *qalli* is written as <قال لي>$_{1}$, <قالي>$_{1}$, and <قاللي>$_{6}$; instead, the constructions *qult^e-l-u/-hum* and *qult^e-l-i* are represented as < قلت له/لهم>$_{5}$ and <قلت للي>$_{1}$. The texts showing a major frequency of compound constructions are the more recent; in [S10] they are 32 vs. 10 and in [S9] 11 vs. 23.

5. Conclusion

From the results of this first phase of my research I have formed the opinion that the two contemporary authors, al-Ḥamīsī and 'Abd al-'Āl, have followed a more restricted set of norms in respect to the past writers, as emerges from a narrower number of variations for each item considered, but at the same time their samples present a higher number of items which, though they could be represented through the standard spelling, are instead written in undoubtedly vernacular forms.

Among the texts considered, Ṣabrī's sample is the one displaying a more coherent and consistent set of choices and that is probably linked to his commitment to a language reform thought.

A special remark is to be devoted to an-Nadīm with his original solutions, at least in the range of my study. He established norms fit for representing the vernacular pronunciation of words containing a glottal stop, corresponding to the standard Arabic phoneme /q/, using the grapheme *hamza*. In fact we find many verbs 'to say' and 'to be' items represented by *hamza* with the three kinds of support or without any, instead of the letter *qāf* which occurs only in a few instances.

Considering all the texts analyzed, other interesting solutions to the literary vernacular writing challenges are found. They could be selected and arranged in a set of norms of how words ought to be written, through a work of planning a writing system, which could help the codification of the Egyptian vernacular as a literary language.

References

AA.VV. 2015. *Encyclopædia Britannica Online*. "Popular art", <http://www.britannica.com/ art/popular-art>.

'Abd al-'Āl, Ġāda. 2008. *'Ayza atgawwez*. Cairo: Dār aš-Šurūq.

Al-Ḫamīsī, Ḫālid. 2006. *Tāksī*. Cairo: Dār aš-Šurūq.

Al-Qa'īd, Yūsuf. 1994. *Laban al-'uṣfūr*. Kotobarabia.com, <www.kotobarabia.com>.

An-Nadīm, 'Abd Allāh. 1985. *Al-Ustāḏ*. Cairo: Dār kutubḫāna li-n-našr wa-t-tawzī'.

At-Tūnisī, Bayram. 1925. *Es-Sayyid we-mrāto fì Maṣr*. Beirut: al-Maktaba al-'aṣriyya.

Avallone, Lucia. 2011. "Autori egiziani degli anni Duemila. Blogosfera, *graphic* e postmoderno: nuovi linguaggi nel panorama letterario arabo", *Kervan 13/14*, 25-46.

'Awaḏ, Luwīs. 2001. *Muḏakkirāt ṭālib bi'ṯa*. Cairo: Dār al Hilāl.

Ḍayf, Šawqī. 2000. "al-'Āmmiyya fuṣḥā muḥarrafa", *Maǧallat Maǧma' al-Luġa al-'Arabiyya bi-l-Qāhira 91*, 31-46.

Daniëls, Helge. 2004. "Linguistic conservatism as the basis for political revolution? The fusha-'ammiya debate in nineteenth-century Ottoman-Aran Middle Eastern society", *Antwerp Papers in Linguistics 106*, 79-92.

El-Enany, Rasheed. 1993. *Naguib Mahfouz: the pursuit of meaning*. London; New York: Routledge.

Eskander, Ramy, & Habash, Nizar, & Rambow, Owen, & Tomeh, Nadi. 2013. "Processing Spontaneous Orthography", *Proceedings of NAACL-HLT 2013*, 585-595.

Hussein, Taha. 1998. *The future of culture in Egypt*. Cairo: The Palm Press.

Imbābī, Fatḥī. 1994. *Marā'ī l-qatl*. Giza: an-Nahr li-n-našr wa-t-tawzī'.

Maḥfūẓ, Naǧīb. 1977. *Ataḥaddaṯ ilaykum*. Beirut: Dār al-'Awda.

Mušarrafa, Muṣṭafā. 1966. *Qanṭara allāḏī kafara*. Cairo: Markaz kutub aš-Šarq al-Awsaṭ.

Rosenbaum, Gabriel. 2004. "Egyptian Arabic as a written language", *Jerusalem Studies in Arabic and Islam 29*, 281-340.

Rosenbaum, Gabriel. 2010. "'I want to write in the colloquial': an example of the language of contemporary Egyptian prose", *Folia Orientalia 47*, 71-97.

Ṣannū', Ya'qūb. 1974. *Ṣuḥuf Abū Naḏḏāra*. Beirut: Dar Sader.

'Uṯmān, Ṣabrī. 1965. *Riḥla fī n-Nīl*. Cairo: Maktabat al-Anglū al-Miṣriyya.

ASSOCIATION
INTERNATIONALE
DE DIALECTOLOGIE
ARABE