

UNIVERSITÉ DE LORRAINE
UFR Arts, Lettres et Langues – Nancy

Master Langues et Sociétés
Parcours Arabe à distance (ERUDI)

Activité intégratrice
Mémoire et Méthodologie de la recherche et du mémoire

Écrire l’arabe égyptien,
le cas d’un Wikipédia “dialectal”

Rédigé par Guy BOURSIER

Direction de mémoire
Laurence DENOZ
Naouel ABDESSEMED
Yannis HARALAMBOUS
2024-2025

Résumé en français

Chaque «dialecte» arabe, ou ‘āmmiyya, est traditionnellement relégué au domaine oral, l’écrit étant l’apanage de la fuṣḥā. Pourtant, des écrits en ‘āmmiyya existent depuis longtemps, notamment en arabe égyptien, à l’origine cantonnés à la transcription de répliques du théâtre, des dialogues dans les romans, des histoires drôles, etc. La création en 2008 d’une édition Wikipédia en arabe égyptien a représenté un palier supplémentaire dans l’incursion de la ‘āmmiyya dans le domaine de l’écrit, à savoir la publication d’articles scientifiques en tous genres. Plus de 15 ans après, elle offre un corpus de plus de 1,6 million d’articles, dont la langue vise à reproduire la variété égyptienne de l’arabe. L’étude des choix graphiques faits par ses multiples rédacteurs révèle une diversité d’emplois des graphes arabes pour transcrire la prononciation égyptienne, mais aussi pour affirmer sa différence d’avec la fuṣḥā. Par une recension des graphes et des graphies utilisés dans ces articles, cette étude établit des statistiques d’emploi de plusieurs stratégies d’écriture pour transcrire les voyelles, notamment en fin de mot, le «coup de glotte», issu ou non du qāf, les alternatives aux dentales /θ/, /ð/ et /ðˤ/, les allophones /ʒ/, /v/ et /p/, la (dé)pharyngalisation, le mot phonique et ses clitiques, etc. Il a été possible de mettre en lumière une créativité graphique spontanée que des recommandations «officielles» essaient d’encadrer dans une tentative originale de définitions de règles orthographiques. Il ressort que l’influence de la tradition fuṣḥā reste souvent prépondérante, mais de multiples innovations graphiques apparaissent, dans un désir de singulariser l’arabe égyptien, encouragé par des processus d’égyptianisation automatisés contribuant à dessiner une norme de l’arabe égyptien écrit.

English abstract

Each Arabic “dialect,” or ‘āmmiyya, is traditionally relegated to the oral domain, writing being the prerogative of the fuṣḥā. However, writings in ‘āmmiyya have existed for a long time, notably in Egyptian Arabic, originally confined to the transcription of theatre lines, dialogues in novels, jokes, etc. The creation in 2008 of a Wikipedia edition in Egyptian Arabic represented a further step in the incursion of ‘āmmiyya into the written domain, namely the publication of scientific articles of all kinds. More than 15 years later, it offers a corpus of more than 1.6 million articles, the language of which aims to reproduce the Egyptian variety of Arabic. The study of the graphic choices made by its multiple editors reveals a diversity of uses of Arabic graphs to transcribe Egyptian pronunciation, but also to assert its difference from fuṣḥā. By a review of the graphs and spellings used in these articles, this study establishes statistics on the use of several writing strategies to transcribe vowels, notably at the end of a word, the “glottal stop”, whether or not derived from qāf, alternatives to the dentals /θ/, /ð/ and /ðˤ/, the allophones /ʒ/, /v/ and /p/, (de)pharyngalization, the phonic word and its clitics, etc. It was possible to highlight a spontaneous graphic creativity that “official” recommendations try to frame in an original attempt at defining spelling rules. It appears that the influence of the fuṣḥā tradition often remains predominant, but multiple graphic innovations crop up, in a desire to singularize Egyptian Arabic, encouraged by automated Egyptianization processes contributing to drawing a standard of written Egyptian Arabic.

Remerciements

Je remercie

- Madame Laurence Denooz, créatrice et responsable de la spécialité Arabe du Master Langues et Sociétés à l'Université de Lorraine pour m'avoir permis de préparer ce diplôme,
- Monsieur Fadi Jaber, co-directeur de ce mémoire lors de la première année de master, pour avoir accepté avec enthousiasme le sujet que je lui ai proposé,
- Madame Naouel Abdessemed, co-directrice de ce mémoire lors de la deuxième année de master, pour avoir spontanément pris la relève de Monsieur Fadi Jaber,
- Monsieur Yannis Haralambous, professeur à l'IMT Atlantique, co-directeur de ce mémoire, pour m'avoir fait bénéficier de ses vastes connaissances en grapholinguistique et m'avoir permis de participer à l'édition 2024 de la conférence *Grapholinguistics in the 21st Century*,
- Madame Marie Varin, Professeure des Universités en linguistique arabe et linguistique de corpus contemporains, détachée à l'Académie Militaire de Saint-Cyr Coëtquidan, pour ses précieux conseils de recherche,
- mes ami.e.s arabophones et égyptiens, notamment Ahmed Aly, pour ses patientes corrections de mon arabe et ses informations inestimables sur la culture égyptienne. Bien sûr, toutes les erreurs restantes sont de mon fait.

Table des matières

Introduction	10
1 L'arabe égyptien et son écriture	14
1.1 Tentative de définition de l'arabe égyptien	14
1.1.1 Dimension géographique	14
1.1.2 Diglossie et dimension socio-culturelle	16
1.1.3 Dimension diachronique	17
1.2 Incursions dans le domaine de l'écrit	18
1.2.1 Écriture de l'arabe égyptien avant la Nahḍa	19
1.2.2 Le théâtre de la Nahḍa	20
1.2.3 Ouverture du champ littéraire	20
1.3 Développement de systèmes d'écriture	23
1.3.1 Adaptation de l'abjad arabe	23
1.3.2 Recours à l'alphabet latin	26
1.4 Conclusion	34
2 Polyvalences de certains graphes arabes	36
2.1 'alif	36
2.1.1 'alif et transcription de la voyelle longue /a:/	37
2.1.2 'alif orthographique	38
2.2 hamza	40
2.3 tā' marbūṭa	41
2.4 wāw et yā'	42
2.4.1 wāw	42
2.4.2 yā'	42
2.5 lām	43
2.5.1 lām 'alif	43
2.5.2 lām al-ta'rif	43
2.6 la gémiation par šadda	44
2.7 Autres assimilations	45
2.7.1 Les verbes de forme viii	45
2.7.2 assimilations non transcrites	46
2.8 Conclusion	46

3	Naissance et enjeux de Wikipédia maşri	47
3.1	Présentation	47
3.2	Débat autour du projet	51
3.3	Élaboration de recommandations orthographiques	55
3.3.1	Suggestions d'écritures en alphabet latin	55
3.3.2	Emploi des graphes arabes pour l'écriture de l'arabe égyptien	56
3.4	Automatisation des choix orthographiques	57
3.4.1	Ghalybot	57
3.4.2	Autres programmes	59
3.5	Conclusion	60
4	Choix orthographiques sur Wikipédia maşri	61
4.1	Fréquence d'utilisation des graphes par rapport à Wikipédia al-‘arabiyya	61
4.2	Voyelles courtes	64
4.2.1	État des lieux	64
4.2.2	Orientations Wikipédia maşri	68
4.2.3	Observations dans le corpus	71
4.3	Voyelles longues	84
4.3.1	État des lieux	84
4.3.2	Orientations Wikipédia maşri	85
4.3.3	Observations dans le corpus	86
4.4	/ʔ/ ou “coup de glotte” issu de la hamza	87
4.4.1	État des lieux	87
4.4.2	Orientations Wikipédia maşri	89
4.4.3	Observations dans le corpus	90
4.5	/ʔ/ ou “coup de glotte” issu de /q/	94
4.5.1	État des lieux	94
4.5.2	Orientations Wikipédia maşri	97
4.5.3	Observations dans le corpus	98
4.6	Fricatives dentales /θ/ et /ð/	100
4.6.1	État des lieux	100
4.6.2	Orientations Wikipédia maşri	102
4.6.3	Observations dans le corpus	103
4.7	Fricatives dentale /ðʕ/, alvéolaire /zʕ/ et occlusive alvéolaire /dʕ/ pharyn- galisées	105
4.7.1	État des lieux	105
4.7.2	Orientations Wikipédia maşri	107
4.7.3	Observations dans le corpus	107
4.8	/g/ consonne occlusive vélaire voisée	109
4.8.1	État des lieux	109
4.8.2	Orientations Wikipédia maşri	110
4.8.3	Observations dans le corpus	111
4.9	[ʒ], [v] et [p] phones allogènes	112
4.9.1	État des lieux	112
4.9.2	Orientations Wikipédia maşri	115
4.9.3	Observations dans le corpus	117

4.10 Assimilation	119
4.10.1 État des lieux	119
4.10.2 Observations dans le corpus	121
4.11 Dépharyngalisation et pharyngalisation dialectale	123
4.11.1 État des lieux	123
4.11.2 Observations dans le corpus	124
4.12 Le mot et ses délimitations	125
4.12.1 État des lieux	125
4.12.2 Orientations Wikipédia maşri	134
4.12.3 Observations dans le corpus	136
4.13 Conclusion	143
Conclusion	144
Bibliographie	147
Pages internet	153

Table des figures

1.1	Préconisations de Ḥalīl ‘Asākir pour transcrire les voyelles des dialectes arabes (Asākir, 1955, p. 185)	24
1.2	Préconisations de Ḥalīl ‘Asākir pour transcrire les consonnes des dialectes arabes (Asākir, 1955, p. 189)	25
1.3	Préconisations de ‘Abd al-‘Azīz Fahmī (Fahmī, 1944, p. 153)	30
1.4	Sūratu l-fātiḥa écrite selon le premier système de ‘Uṭmān Ṣabrī (Al-Sāwī, 2017, p. 16)	31
1.5	‘Āyatu l-kursī écrite selon le deuxième système de ‘Uṭmān Ṣabrī (Al-Sāwī, 2017, p. 19)	31
1.6	Système d’écriture proposé par Ṣa‘īd ‘Aql (Al-Sāwī, 2017, p. 21)	32

Liste des tableaux

1.1	Systèmes de transcription de l'arabe égyptien (fin xix ^e , début xx ^e siècles)	28
1.2	Systèmes de transcription de l'arabe égyptien (fin xx ^e , début xxi ^e siècles)	33
2.1	Lettres et phonèmes arabes en relation isomorphe	36
3.1	Nombre annuel d'articles créés sur Wikipédia mašri	48
3.2	Comparaison de différentes éditions de Wikipédia	50
3.3	Arguments sur la concurrence avec Wikipédia arabe	51
3.4	Arguments sur le statut de l'arabe égyptien	52
3.5	Arguments sur la richesse de l'arabe égyptien	53
3.6	Développement d'un article de l'anglais vers la fuṣḥā puis la 'āmmiyya .	58
4.1	Sous-représentation des graphes arz par rapport à ar	62
4.2	Sur-représentation des graphes arz par rapport à ar	63
4.3	Fréquence des diacritiques sur arz et ar	72
4.4	Fréquence de ى et ٰ sur arz et ar	72
4.5	Fréquence de َ et ِ sur arz et ar	73
4.6	Graphies des finales de /kunt-/ et /ʔint-/	76
4.7	/i/ final des prépositions	77
4.8	Prépositions /li/ et /bi/ avec pronoms clitiques de 3 ^o personne	77
4.9	/u/ final transcrit par ٰ ou ِ	78
4.10	/u/ final transcrit par ِ ou ِ	78
4.11	/a/ final transcrit par ِ, ِ ou ِ	79
4.12	/a/ final transcrit par ِ, ِ, ِ ou ِ	80
4.13	/a/ final transcrit par ِ, ِ ou ِ	81
4.14	Fréquence de la tā' marbūṭa par rapport à hā'	82
4.15	Fréquence de la faṭḥatān	83
4.16	Fréquence des graphes yā' ِ, wāw ِ et 'alif ِ sur arz et ar	86
4.17	Fréquence de la hamza et madda sur arz et ar	91
4.18	/ʔa/ ou /ʔa/ initial transcrit par ِ ou ِ	91
4.19	/ʔa:/ ou /ʔa:/ initial transcrit par ِ ou ِ	92
4.20	Emploi ou omission de hamza initiale (Ghalybot)	92
4.21	Emploi ou omission de hamza initiale (hors Ghalybot)	93
4.22	/ʔu/ initial transcrit par ِ ou ِ	93
4.23	/ʔi/ initial transcrit par ِ, ِ ou ِ	94

4.24	Hamza médiane conservée ou non	94
4.25	/ji/ médian transcrit par ي ou ء	95
4.26	/ʔu:/ médian transcrit par hamza sur différents supports	95
4.27	Hamza finale conservée ou non	96
4.28	Passage de /q/ en fuṣḥā à /ʔ/ en ‘āmmiyya	96
4.29	Paires minimales distinguées par /q/ ou /ʔ/	97
4.30	Fréquence du graphe ق sur arz et ar	98
4.31	/ʔ/ issu de /q/ transcrit par ق ou ا	100
4.32	Devenir de /θ/ et [ð] fuṣḥā en arabe égyptien	101
4.33	Transformation de [θ] et [ð] dans <i>A Dictionary of Egyptian Arabic</i>	101
4.34	Devenir différent de /θ/ et /ð/ pour une même racine	101
4.35	Traitement du graphe ث par Ghalybot	103
4.36	Traitement du graphe ذ par Ghalybot	103
4.37	Fréquence des graphes ث et ذ sur arz et ar	104
4.38	/z/ transcrit par ذ ou ز	104
4.39	/s/ transcrit par ث ou س	104
4.40	/d/ transcrit par ذ ou د	105
4.41	/t/ transcrit par ث et ت	105
4.42	Passage de /ðˢ/ en fuṣḥā à /zˢ/ ou /dˢ/ en ‘āmmiyya	106
4.43	Passage de /dˢ/ en fuṣḥā à /zˢ/ en ‘āmmiyya	106
4.44	Traitement du graphe ظ par Ghalybot	107
4.45	Fréquence des graphes ظ et ض sur arz et ar	108
4.46	/dˢ/ transcrit par ظ ou ض	108
4.47	/zˢ/ transcrit par ظ, ض ou ز	109
4.48	Fréquence des graphes ج et غ sur arz et ar	111
4.49	/g/ ou /ɣ/ transcrits par ج ou غ	112
4.50	Fréquence des graphes alternatifs چ, ف et پ sur arz et ar	117
4.51	[ʒ] transcrit par چ ou ج	118
4.52	[v] transcrit par ف ou ف	118
4.53	[p] transcrit par پ ou ب	119
4.54	Assimilation du lām al-ta‘rīf spécifique à l’arabe égyptien	120
4.55	/m/ transcrit par م ou ن	122
4.56	/n/ transcrit par م et ن	122
4.57	Pharyngalisation d’emprunts étrangers marquée ou non	125
4.58	Pronoms personnels clitiques en arabe égyptien	126
4.59	Négation du verbe en arabe égyptien	127
4.60	Négation des prépositions avec clitiques en arabe égyptien	127
4.61	Nombres de onze à dix-neuf	128
4.62	Les centaines	129
4.63	Les milliers	129
4.64	Particules du futur ح et ها séparées du verbe	136
4.65	Particules du futur ح et ها liées au verbe	137
4.66	Graphies de la particule négative /ma ... ʃ/	137
4.67	Graphies de la préposition abrégée /ʕa-l...	138

4.68	Graphies de la préposition /li/ + pronom personnel	140
4.69	Graphies des adjectifs numéraux de 11 à 19	141
4.70	Graphies des centaines	142
4.71	Graphies des milliers	142

Introduction

Je prouveray en barbe de je ne scay
quels centonifiques botteleurs de
matières cent et cent fois gabelées,
rappetasseurs de vieilles ferrailles
latines, revendeurs de vieux mots latins
tous moisis et incertains, que nostre
langue vulgaire n'est tant vile, tant
inepte, tant indigente et à mespriser
qu'ils l'estiment.

François Rabelais
Le Cinquiesme et dernier livre

L'arabe littéral, issu de l'arabe classique, est une langue écrite par excellence. Elle s'apprend essentiellement à l'école et représente une norme qui s'appuie sur un corpus de textes écrits. Sa pratique orale est définie par le respect de règles déduites de ce corpus écrit. L'arabe égyptien, comme les autres régiolectes classés parmi les dialectes de l'arabe, est au contraire tout d'abord une pratique orale spontanée, acquise dans l'intimité du cercle familial, puis au rythme des interactions sociales. Son écriture a traditionnellement été considérée comme marginale et déviante par rapport à la norme que représente l'arabe littéral. De tous ces régiolectes arabes, l'arabe égyptien présente cependant un des plus larges corpus de textes écrits, et une des pratiques scripturales les plus affranchies de la règle d'invisibilisation induite par la suprématie officielle de l'arabe littéral.

Études de l'écriture de l'arabe égyptien

Ce phénomène d'écriture de l'arabe égyptien a ainsi été l'objet d'études portant notamment sur les choix d'emploi des graphes arabes pour écrire cette variété de l'arabe.

En 1981, Refaat El-Farnawany, un Égyptien, a soutenu à Nuremberg sa thèse de doctorat intitulée *Ägyptisch-arabisch als geschriebene Sprache : Probleme der Verschriftung einer Umgangssprache* 'L'arabe égyptien comme langue écrite : les problèmes d'écriture d'un dialecte'. Après un exposé de la phonologie de l'arabe égyptien, il présente la façon dont la langue est écrite, graphe arabe par graphe arabe, puis les traces écrites de phénomènes comme l'assimilation et la métathèse, la séparation ou l'assemblage entre morphèmes et enfin le traitement des mots d'origine étrangère. Il puise ses exemples dans

un corpus de plus de 110 œuvres dont la date de publication s'étend des années 1920 aux années 1970 (Farnawany, 1981).

En 1983, l'Institut Français d'Archéologie Orientale du Caire publie *L'Égyptien tel qu'on l'écrit, Glossaire établi d'après un choix d'œuvres littéraires égyptiennes contemporaines* de Charles Vial, qui décrit plusieurs phénomènes orthographiques de l'arabe égyptien dans son introduction. Il s'appuie sur un corpus de 80 ouvrages (dont trois seulement en commun avec le précédent) publiés des années 1940 au début des années 1980 (Vial, 1983, p. X-XII).

En 2004, Gabriel M. Rosenbaum publie dans *Jerusalem studies in Arabic and Islam* un article intitulé "Egyptian Arabic as a written language" (L'arabe égyptien en tant que langue écrite) où il passe en revue de nombreuses manières de mettre à l'écrit l'arabe égyptien. Il s'intéresse à l'emploi des consonnes, des voyelles, aux modifications du concept d'unité de mot, à la tension entre convention orthographique et prononciation. Son corpus est constitué de plus de 180 œuvres (22 en commun avec Farnawany et 3 avec Vial) publiées des années 1950 aux années 2000 (Rosenbaum, 2004).

En 2016, paraît dans *Proceedings of the 11th International Conference of AIDA*, un article de Lucia Avallone (Université de Bergame) intitulé "Spelling variants in written Egyptian Arabic, a study on literary texts" (Variantes orthographiques dans l'arabe égyptien écrit, étude sur des textes littéraires). L'auteure étudie 10 échantillons de 20 à 40 pages d'œuvres rédigées en arabe égyptien et publiées entre 1878 et 2008, et décrit pour chacun des échantillons les choix orthographiques faits par les auteurs (Avallone, 2015).

Toutes ces études sont autant de recensions de multiples façons d'utiliser les graphes arabes pour écrire l'arabe égyptien, et démontrent toutes l'absence d'une norme orthographique et la fréquence des variations d'écriture d'un même mot, parfois par un même auteur. Les corpus étudiés sont essentiellement constitués d'œuvres littéraires. Le théâtre y occupe une place prépondérante, sa dimension orale le rendant propice aux incursions de la langue parlée dans ses textes. Nombreux sont aussi les récits à la première personne, ou écrits selon la technique du "courant de conscience" transcrivant le processus de pensée d'un individu. Là aussi, l'adoption d'une langue intérieure est propice à l'emploi de l'arabe de tous les jours.

Wikipédia en arabe égyptien

La création en 2008 d'une édition de Wikipédia en arabe égyptien fut une première pour un dialecte arabe, ce qui confirme le dynamisme des promoteurs de ce régiolecte. Il s'agit cependant d'un tout autre projet que la plupart des publications habituelles en arabe égyptien. En effet, selon la dichotomie établie par Benveniste (1966, p. 237), celles-ci relèvent essentiellement du plan d'énonciation du discours, alors qu'une encyclopédie se situe sur celui de l'histoire, avec effacement des première et deuxième personnes, ainsi que, théoriquement, du temps présent, sauf le "présent de définition", et donc de toute subjectivité. De plus, la langue doit ici être mise au service d'un projet encyclopédique, avec des attentes de rigueur scientifique et de précision terminologique peu habituelles dans son usage quotidien. Le contenu de cette encyclopédie en arabe égyptien représente donc un corpus apte à mettre en évidence de nouvelles stratégies d'adaptation à l'écrit de cette langue.

Wikipédia en arabe égyptien a déjà fait l'objet de plusieurs études, la première datant de 2010, soit à peine deux ans après le lancement de cette édition. Il s'agit d'un article de 35 pages intitulé "The beginnings of Wikipedia masry" (Les débuts de Wikipédia masri), écrit par Ivan Panovic et publié dans la revue *al-Logha*. Il s'agit d'une étude socio-linguistique du jeune projet qui met déjà en lumière sa dimension nationaliste et les tentatives de différenciation de l'arabe égyptien de l'arabe littéral (Panović, 2010).

En 2023, l'Université de Varsovie publie la thèse de Zuzanna Nabalssi-Maselbas intitulée *Egyptian Linguistic Separatism, A Study in Wikipedia Masri* (Séparatisme linguistique égyptien, une étude de Wikipédia Masri). Elle y présente longuement les mouvements nationalistes égyptiens, dont le "pharaonisme", qui ont précédé l'apparition de Wikipédia en arabe égyptien, et la manière dont ce projet s'inscrit dans cette dynamique. Elle qualifie cependant les types d'écriture non orthodoxes de "fautes" sans véritablement envisager leur potentiel d'innovation (Nabalssi-Maselbas, 2023).

Présentation de cette étude

La présente étude s'appuiera donc sur les recherches faites précédemment dans le domaine.

Le **corpus** sera constitué des pages de l'édition en arabe égyptien de Wikipédia, désignée par le terme مصرى "masri"¹ dans le menu linguistique de l'encyclopédie en ligne, et dont l'adresse web est <https://arz.wikipedia.org>.

Ma **problématique** est la suivante : Comment les rédacteurs des articles de Wikipédia en arabe égyptien résolvent-ils le problème de la mise à l'écrit d'une langue essentiellement parlée et dont les règles d'orthographe ne font l'objet d'aucune norme établie ?

Mon **hypothèse de recherche** est que les rédacteurs bénévoles ont élaboré par consensus interne un certain nombre de règles de base, mais que ces règles unificatrices ne sont pas systématiquement suivies, en raison de l'influence d'autres formes de la langue, comme l'arabe littéral, les régionalismes, la confrontation à des langues étrangères, ...

Ma **méthodologie** sera la suivante : Après un passage en revue des différentes tentatives de réalisation d'un système d'écriture de l'arabe égyptien, en script arabe ou autre, j'étudierai les préconisations faites aux rédacteurs d'article en arabe égyptien, en script arabe, au sein de l'encyclopédie Wikipédia. J'analyserai les choix faits et leur pertinence par rapport, d'une part, à l'utilisation canonique du script arabe pour l'arabe littéral et, d'autre part, aux particularismes, notamment phonologiques, de l'arabe égyptien. J'étudierai ensuite les débats auxquels l'élaboration de ces choix aura donné lieu afin de mettre en évidence les différentes stratégies de passage à l'écrit envisagées. Je relèverai dans le corpus des exemples de respect et de non-respect de ces préconisations de la communauté pour ce qui concerne l'écriture et l'orthographe. En confrontant préconisations et pratiques, mon objectif est de dessiner les lignes de tension qui émergent de cette tentative d'écriture de l'arabe égyptien à grande échelle dans un domaine inhabituel.

Pour cela, cette étude s'attachera à montrer l'impact du principe de pluralité des auteurs sur la langue produite. La dimension numérique du corpus permettra un recensement

¹Pour cet adjectif signifiant "égyptien", je privilégierai la prononciation égyptienne [mas^ʕri] (Badawi et Hinds, 1986, p. 826) (plutôt que la classique [mis^ʕri:] (Wehr et Cowan, 1979, p. 1070), incité par l'absence des deux points sous le yā' final, trait caractéristique de l'écriture égyptienne).

automatisé des formes jugées pertinentes pour mettre en évidence leur poids statistique et ainsi contribuer à dessiner un état de l'arabe égyptien écrit. J'essaierai aussi d'apporter des réponses aux questions que pose Florian Coulmas dans *Writing systems, An introduction to their linguistic analysis* (Systèmes d'écriture, introduction à leur analyse linguistique) : "What happens when a language is written down, (1) in terms of linguistic description, and (2) in terms of linguistic evolution ?" (Que se passe-t-il lorsqu'une langue est mise à l'écrit, (1) en termes de description linguistique et (2) en termes d'évolution linguistique ?) (Coulmas, 2003, p. 11).

Je considérerai le graphème selon son rôle décrit par Catach (1979, p. 26) en tant que "signifiant (forme écrite) renvoyant à un signifiant forme orale", l'écriture de l'arabe égyptien se faisant essentiellement par le recours aux graphes du script arabe auxquels correspondent les phonèmes de la fuṣḥā. J'organiserai donc mon étude en parcourant la liste des phonèmes de l'arabe égyptien et en décrivant les différentes manières de les transcrire sur Wikipédia maṣri.

Point terminologique : J'utiliserai les termes "arabe égyptien" ou "‘āmmiyya" pour faire référence aux langues à laquelle l'ISO 639-3 attribue le code arz (*Egyptian Arabic*) et le code aec (*Saidi Arabic*). Si nécessaire, pour distinguer les deux régiolectes, je préciserai "arabe du Caire" d'une part et "arabe ṣa‘īdi" d'autre part. Je réserverai le terme "dialectal" à la référence générale aux différents géolectes de l'arabe (formes d'arabe parlé qui diffèrent selon les pays ou zones géographiques). J'utiliserai les termes "arabe littéral" ou "fuṣḥā" pour faire référence à la langue à laquelle l'ISO 639-3 attribue le code arb (*Standard Arabic*), et "arabe" pour la macro-langue (code ara).

Translittération et transcription : Le système de translittération de la revue *Arabica* sera utilisé pour rendre les écrits arabes tels que les titres d'œuvres, les noms des auteurs, ... J'utiliserai l'alphabet phonétique international (API) pour la transcription phonologique. Je reproduirai la graphie arabe quand je souhaiterai montrer les phénomènes qui y ont cours.

Chapitre 1

L'arabe égyptien et son écriture

1.1 Tentative de définition de l'arabe égyptien

Définir ce qu'est l'arabe égyptien requiert d'étudier cette réalité linguistique sous différents angles.

1.1.1 Dimension géographique

Parler d'« arabe égyptien » implique une définition spatiale de ce phénomène linguistique. Cette appellation part d'une langue, l'arabe, une des langues afro-asiatiques (appelées aussi langues chamito-sémitiques), relevant plus précisément du groupe des langues sémitiques¹, dont l'origine géographique est l'ouest de la péninsule arabique, et dont l'emploi s'est développé tout au long de l'Histoire sur plusieurs continents. L'arabe compte actuellement près d'un demi-milliard de locuteurs et est pratiqué couramment dans des dizaines de pays, dont 22 sont membres de la Ligue arabe.

Au terme « arabe », est adjointe l'épithète « égyptien » pour faire référence à une variété spécifique définie géographiquement : celle pratiquée par les habitants de l'Égypte, ce pays centré sur la partie nord de la vallée du Nil, dont le delta de ce fleuve. La géographie physique de la république arabe d'Égypte concentre la plupart de ses plus de 100 millions d'habitants dans cette zone fertile, même si son territoire inclut de vastes portions du Sahara à l'est et surtout à l'ouest. Les variétés d'arabe qui sont pratiquées sur les bords du Nil sont donc relativement distinctes de leurs voisines occidentales (arabe libyen) et orientales (arabe bédouin du désert oriental, arabe levantin et arabe du Hedjaz, par delà la Mer Rouge), tout en s'inscrivant dans un continuum nord-sud qui s'étend jusqu'aux variétés de l'arabe soudanais au sud.

¹Outre le groupe des langues sémitiques, parlées au Moyen-Orient, dans la Corne de l'Afrique, en Afrique du Nord et dans la Méditerranée, y sont notamment rattachés celui des langues berbères, parlées en Afrique du Nord et au Sahara, celui des langues couchitiques, parlées dans la Corne de l'Afrique et le long des côtes de la mer Rouge, du Soudan au sud-est égyptien, celui de l'égyptien ancien et son descendant le copte (aujourd'hui éteint et conservé seulement comme langue liturgique), jadis parlés en Égypte, celui des langues tchadiques, parlées en Afrique sahélienne, du Niger à la Centrafrique.

On dénombre environ 35 langues dont l'appellation contient le terme «*Arabic*» dans l'ISO 639-3, une norme internationale qui définit des codes à trois lettres pour identifier les langues² (ISO, 2024).

L'ISO 639-3 propose le découpage suivant des langues arabes en Égypte :

- arz *Egyptian Arabic* = arabe égyptien ;
- aec *Saidi Arabic* = arabe ša'īdi = arabe de Haute-Égypte ;
- avl *Eastern Egyptian Bedawi Arabic* = arabe bédouin du désert oriental d'Égypte ;
- ayl *Libyan Arabic – Western Egyptian Bedawi Arabic* = arabe bédouin du désert occidental d'Égypte.

On remarque qu'en utilisant l'appellation «arabe égyptien» pour une de ces langues une certaine incohérence terminologique se fait jour. Ceci met en évidence la difficulté de nommer la langue parlée par la majorité des Égyptiens, en particulier les Cairotes, sans exclure les autres régiolectes du pays, notamment l'«arabe ša'īdi» l'autre variété d'arabe largement utilisée en Égypte. Cette taxonomie n'échappe pas à un certain «ethnocentrisme» et laisse entrevoir la moindre estime dans laquelle sont tenues traditionnellement, en particulier au Caire, la culture de Haute-Égypte et les cultures bédouines extérieures à la vallée du Nil et à son delta.

Dans leur *Dialektatlas von Ägypten* 'Atlas des dialectes de l'Égypte' (1985), Peter Behnstedt et Manfred Woidich traitent d'abord à part les régiolectes des bédouins qui vivent dans les marges (Sinaï, déserts occidentaux et orientaux) et ne sont pas encore complètement sédentarisés. Ils divisent ensuite les régiolectes de la population rurale sédentaire actuelle en trois zones géographiques principales : le delta du Nil, la vallée du Nil et les oasis occidentales, zones elles-mêmes divisées en sous-groupes. Le delta du Nil comprend les régiolectes occidentaux, centraux (dont est proche l'arabe du Caire), du nord-est et orientaux. La vallée du Nil comprend les régiolectes de Moyenne-Égypte et de Haute-Égypte. Les oasis occidentales comprennent les régiolectes des oasis al-Bahariyya, al-Farāfra, al-Dāhla, al-Ḥārġa (Behnstedt et Woidich, 1987).

L'appellation «arabe égyptien» fait cependant le plus souvent référence à la variété d'arabe parlée dans la capitale égyptienne. El-Said Muhammad Badawi, (1929-2014) et Martin Hinds (1941-1988) indiquent ainsi notamment dans l'introduction à *A Dictionary of Egyptian Arabic : Arabic-English* (Badawi et Hinds, 1986, p. XI) que celui-ci se concentre sur la langue utilisée au Caire, le cadre phonologique étant celui du Caire et du Delta. Ils différencient cette variété de celle de Haute-Égypte, dont la phonologie est sensiblement différente. Ils font aussi allusion à une autre variété parlée à Alexandrie. Les manuels d'apprentissage, lexiques ou grammaires de l'arabe égyptien, destinés aux allochtones, précisent aussi généralement dans leur titre que la langue qu'ils décrivent est celle pratiquée au Caire, avec parfois la référence «parler du Caire» (Jomier et Khouzam,

²L'ISO 639 est une norme de l'Organisation internationale de normalisation (ISO) qui définit des codes pour la représentation des noms de langues. Les normes ISO 639-1 et ISO 639-2 se concentraient sur les langues majeures (dont l'arabe). L'ISO 639-3 les étend afin de couvrir toutes les langues naturelles connues (dont les "dialectes" comme l'arabe égyptien), en s'appuyant principalement sur les codes linguistiques utilisés dans la base de donnée *Ethnologue* publiée par SIL International.

1973), «arabe égyptien du Caire» (Boutros, 1993) en français ou «*spoken Arabic of Cairo*» (Stephens et Maurice, 2004) en anglais. L'arabe du Caire est essentiellement un dialecte du delta central mais partage un certain nombre de caractéristiques avec tous les autres dialectes de son voisinage (Manfred Woidich, 1994, p. 306). L'attraction qu'exerce la capitale égyptienne implique que de nombreux locuteurs de variétés géographiques éloignées ont inéluctablement apporté leurs particularismes. L'arabe du Caire a été décrit comme une langue hybride qui a été plus ou moins façonnée par l'afflux massif de migrants ruraux dans la seconde moitié du xix^e siècle (Youssef, 2013, p. 19). Cette variété d'arabe ne peut donc être considérée comme une langue «pure», isolée.

1.1.2 Diglossie et dimension socio-culturelle

Dans son article de 1959 intitulé «Diglossia», le socio-linguiste américain Charles A. Ferguson fait de l'arabe un exemple parmi quatre³ pour décrire le phénomène de diglossie où deux variétés linguistiques coexistent sur un même territoire, en l'occurrence l'arabe littéral (fuṣḥā) et l'arabe dialectal (‘āmmiyya) (Ferguson, 1959a). Selon les locuteurs des deux variétés, le premier est considéré comme supérieur, requérant un apprentissage scolaire, et apte à être employé dans des contextes formels religieux, littéraires, scientifiques, ... Le second s'acquiert naturellement comme langue maternelle⁴ et est utilisé dans les échanges du quotidien, essentiellement à l'oral.

Cette description dichotomique a été nuancée par El-Said Muhammad Badawi, dans son livre *Mustawayāt al-‘arabiyya al-mu‘āṣira fī miṣr* 'Niveaux de l'arabe contemporain en Égypte' (Badawi, 1973), nuance reprise dans l'introduction à *A Dictionary of Egyptian Arabic : Arabic-English* (Badawi et Hinds, 1986, p. VIII-X). Il y décrit cinq niveaux d'arabe en Égypte, fuṣḥā et ‘āmmiyya se répartissant autant de zones d'un système au sein duquel les Égyptiens communiquent entre eux dans des domaines très divers.⁵

On peut ainsi considérer que la différence entre les deux variétés linguistiques clairement distinguées correspond à un continuum que le locuteur parcourt en fonction de sa capacité à maîtriser tel ou tel niveau et de ses besoins communicationnels du moment. Il y a donc une pluralité d'emplois de l'arabe égyptien correspondant à la pluralité des situations d'élocution.

La relation entre les deux pôles de ce continuum n'est cependant pas neutre, Ferguson décrivant l'un comme haut (fuṣḥā) et l'autre comme bas (‘āmmiyya). Les appellations elles-mêmes illustrent cette hiérarchisation, fuṣḥā signifiant «la plus éloquente», et ‘āmmiyya «la générale», réservée au commun du peuple.

Cette perception a toutefois été critiquée. William Willcocks (1852-1932), l'ingénieur en irrigation britannique qui a supervisé l'édification du barrage bas d'Assouan et de celui d'Assiout, publia en 1893, dans le magazine *Mağallat al-Azhar* un article intitulé «Limā lam tūğad quwwat al-iḥtirā‘ ladā l-miṣriyyīn ? » "Pourquoi les Égyptiens n'ont-ils aucune force créatrice ?". La réponse à cette question était, pour lui, la situation de diglossie dans

³Les trois autres sont le grec moderne, le suisse allemand et le créole haïtien.

⁴L'arabe égyptien s'acquiert aussi par apprentissage formel à partir de livres, en tant que langue étrangère, notamment par les étrangers vivant en Égypte, d'où la publication de manuels à cet effet.

⁵Les cinq niveaux définis par Badawi sont ‘āmmiyyatu l-‘ummiyyīn (dialecte des illettrés), ‘āmmiyyatu l-mutanawwarīn (dialecte des éclairés), ‘āmmiyyatu l-muṭaqqafīn (dialecte des cultivés), fuṣḥā l-‘aṣr (arabe littéral contemporain), fuṣḥā l-turāṭ (arabe littéral de patrimoine).

laquelle vivaient les Égyptiens. Il les appelait donc à «écrire la langue égyptienne et à arrêter de perdre leur temps à apprendre l'arabe [littéral] qui [...] les prive de leur énergie [intellectuelle]» (Hanna, 2019, p. 368). Cet article a suscité de nombreuses controverses parmi les intellectuels égyptiens de l'époque, et il est encore souvent cité pour dénoncer un désir colonial de nuire à la nation arabe ou à la communauté islamique en promouvant la division par la légitimation des dialectes dans le monde arabe.

D'autres Occidentaux, comme John Selden Willmore (1856-1931), un des juges de la cour d'appel du Caire, considèrent ainsi que la situation de diglossie que connaît la population égyptienne est un obstacle à son développement (Willmore, 1901, p. xii). L'analphabétisme, qui était alors prépondérant dans la population générale, établissait ainsi une barrière entre la majorité des gens du peuple et une élite lettrée, l'apprentissage de l'écrit nécessitant en effet l'acquisition d'une langue différente de celle parlée. Ces considérations étaient souvent mues par le désir sincère de faire bénéficier l'Égypte des progrès scientifiques récents, même si elles étaient aussi teintées de mépris et de morgue, sentiments générés par le statut de colons, ou équivalent, de ces Occidentaux.

À l'occasion de la glorification d'un certain nationalisme égyptien au début du xx^e siècle, connu sous l'appellation de Pharaonisme, des figures égyptiennes expriment des idées similaires. On attribue ainsi à Qāsim Amīn (1865-1908), juriste égyptien, réformateur islamique, l'un des fondateurs du mouvement national égyptien ainsi que de l'Université du Caire, la phrase suivante : إن الأوروبي يقرأ لكي يفهم أما نحن نفهم لكي نقرأ «L'Européen lit pour comprendre, mais nous, nous comprenons pour lire» (Mūsā, 1926, p. 1073). Aḥmad Luṭfi Sayyid (1872 - 1963), journaliste et homme politique égyptien, appelle en 1913 dans son journal al-Ġarīda à donner plus de place à la 'āmmiyya dans l'écriture en fuṣḥā (Diem, 2006, p. 133-134).

Salāma Mūsā (1887-1958), écrivain, journaliste et réformateur égyptien, publie en 1926 dans le magazine *al-Hilāl* un article intitulé «al-luġa al-fuṣḥā w al-luġa al-'āmmiyya» où il appelle à «égyptianiser» l'arabe littéral (Mūsā, 1926, p. 1075). Ces orientations sont toutefois condamnées, notamment par Muṣṭafā Ṣādiq al-Rāfi'ī (1880-1937), poète égyptien, qui défend l'inviolabilité de la fuṣḥā en tant que langue du Coran et moyen de transmission du patrimoine culturel arabe (Diem, 2006, p. 134-135).

La situation diglossique de la langue arabe en Égypte est donc le théâtre de tensions entre les défenseurs des différentes variétés, qu'elles soient perçues comme «hautes» ou comme «basses».

1.1.3 Dimension diachronique

Une troisième dimension nécessaire pour décrire l'arabe égyptien est celle de l'Histoire. Au sein même de la langue arabe, une situation de diglossie est postulée dès la période préislamique. Selon Pierre Larcher «la plupart des arabisants font l'hypothèse que l'arabe connaissait une situation de diglossie dès avant l'islam (ou, s'il ne la connaissait pas, que celle-ci est apparue en moins d'un siècle après l'islam)» (Larcher, 2001, p. 604-605). Dans sa lecture de l'article «'Arabiyya» de l'*Encyclopédie de l'islam*, il décèle un «modèle diglossique» :

L'histoire de l'arabe [est] vue comme une suite de diglossies, où l'une chasse l'autre.
À la diglossie préislamique (dialectes anciens vs koinè poétique), succède, la koinè

poétique et l'arabe coranique étant la base de l'arabe classique, une diglossie postislamique (arabe classique vs koinè militaro-urbaine), à laquelle succède, la koinè militaro-urbaine étant la base des dialectes modernes et l'arabe classique celle de l'arabe moderne, la diglossie contemporaine (arabe (standard) moderne vs dialectes modernes) ... (Larcher, 2001, p. 589)

Dans un deuxième article de 1959, intitulé «The Arabic Koine», Charles A. Ferguson postule que la plupart des dialectes arabes modernes sont issus d'une forme d'arabe, qu'il nomme koinè, qui n'était identique à aucun des dialectes antérieurs et qui différait à bien des égards de l'arabe classique mais était utilisée côte à côte avec l'arabe classique au cours des premiers siècles de l'ère musulmane. Pour appuyer sa démonstration, il décrit quatorze caractéristiques sur lesquelles les dialectes modernes s'accordent en contraste avec l'arabe classique (Ferguson, 1959b).

Pour l'Égypte, sa conquête par l'armée arabe dirigée par 'Amr b. al-ʿĀṣ autour de 640 EC a imposé la langue des vainqueurs comme langue administrative dans le delta et la vallée du Nil. Elle a ainsi pris la place de la koinè grecque usitée au sein de l'appareil administratif de l'empire byzantin, alors que la majeure partie de la population a continué pendant plusieurs siècles à parler la langue copte. Une nouvelle situation de bilinguisme est donc apparue, avec la coexistence de l'arabe pratiqué par l'élite gouvernante (en lieu et place du grec) et du copte conservé par les populations locales.

Sous l'effet combiné de l'immigration à grande échelle d'Arabes de la péninsule et de la conversion parfois contrainte d'une grande partie des Coptes à l'islam à la fin du viii^e et au ix^e siècle, l'arabe est devenu la principale langue de communication. Au cours des x^e et xi^e siècles, en quelques générations, le copte a disparu en tant que langue parlée (Rubenson, 1996, § 5).

Il semble que l'influence du copte sur l'arabe en Égypte ait été assez limitée. Bien qu'il existe de nombreux emprunts coptes dans l'arabe égyptien, en particulier en Haute-Égypte, les Coptes n'ont pas développé leur propre dialecte arabe, ni par écrit ni, pour autant que l'on puisse en juger d'après les preuves de l'époque, dans la langue vernaculaire. L'influence de cette langue ne subsiste donc que sous la forme d'un substrat relativement limité (ibid., § 33).

La langue parlée en Égypte a ainsi évolué au fil des siècles, comme les écrits qui subsistent l'attestent pour aboutir, en parallèle avec l'islamisation de la société, à l'arabe égyptien contemporain, incorporant de nombreux termes turcs, puis européens (italiens, français, anglais, ...). Ce parler continue de se transformer, l'emprunt à l'anglais devenant de plus en plus commun (Rosenbaum, 2014).

1.2 Incursions dans le domaine de l'écrit

L'arabe égyptien a traditionnellement été relégué au domaine de l'oral et de l'informel, celui de l'écrit et de l'officiel étant accaparé par la fuṣḥā. Toutefois, la frontière entre les deux n'est pas si étanche que cela.

1.2.1 Écriture de l'arabe égyptien avant la Nahḍa

Des traces de dialectes sont décelables dans les écrits depuis des siècles, le système d'écriture recourant aux graphes arabes⁶ utilisés selon le principe de l'abjad, c'est-à-dire un système d'écriture privilégiant la notation des consonnes et voyelles longues, au détriment des voyelles courtes.

Si on prend en compte les types d'écritures mélangeant langue littéraire et dialectes, il y a trace de ces derniers à toutes les époques. On trouve ainsi d'une part « dans des textes à caractère principalement classique, des éléments de vocabulaire, de morphologie ou de construction rappelant la langue parlée de différentes époques » et d'autre part « certains textes écrits en dialectal [...] rarement dénués de traits appartenant à la langue écrite, littéraire. » (Doss, 1996, § 2). Ceci a amené à postuler l'émergence d'un « moyen arabe », que Madiha Doss définit comme « tout texte dont la vocation est d'être écrit en classique mais où apparaissent des traits de la langue parlée ou, du moins, des aspects du classique sous-normé. » (ibid., § 7).

Manuel Sartori de son côté montre comment des manuscrits de grammaire arabe médiévaux ne sont eux-mêmes pas exempts de traits dialectaux et relèvent, malgré les attentes qu'ils suscitent, d'un entre-deux aux confins de la fuṣḥā et de la 'āmmiyya (Sartori, 2014).

En ce qui concerne les œuvres littéraires, pour ce qui est de l'arabe égyptien, on peut citer le recueil d'anecdotes satiriques connu sous le nom de *Kitāb al-fāṣūṣ fi Ahkām Qarā-qūṣ*, ou « Livre sur la stupidité des jugements de Qaraqush », commencé par Ibn Mammati (1149-1209) (Dowaidar, 2020), ainsi que les pièces d'Ibn Dānyāl (m. 1310) pour le théâtre d'ombres (ḥayāl al-zill) (Danyal, 2015). Ces œuvres médiévales furent parmi les premières à donner au dialecte une dimension de langue littéraire. La mise par écrit d'anecdotes transmises oralement ou de dialogues de théâtre apparaît comme une porte d'entrée naturelle de l'arabe égyptien dans ce domaine prestigieux qu'est la littérature.

Ultérieurement, certains textes en prose sont apparus à l'époque ottomane, mais aucun long texte en prose n'a été trouvé avant le xvii^e siècle. C'est l'époque où Yūsuf al-Širbīnī écrit *Hazz al-Quḥūf bi Šarḥ Qasīd Abī Šādūf*, une compilation qui contient notamment un long poème de zağal⁷, en arabe égyptien avec un commentaire en prose également en 'āmmiyya (Shirbīnī et Davies, 2005). Le livre *Waqāyi' Miṣr al-Qāhira al-Maḥrūsa*, ou chroniques de Damurdashi (1686-1756), qui relate des événements advenus en Égypte durant la seconde moitié du xvii^e siècle (Damardāsh, 2002), est « un des rares ensembles d'écrits portant si clairement des traces d'oralité » (Doss, 1996, § 44).

Il existe aussi un continuum de textes qui va des documents écrits par des personnes dont le métier n'est pas d'écrire mais qui rédigent, pour les besoins de la vie pratique, des lettres ou des requêtes, jusqu'aux textes écrits par des personnes qui ont reçu une éducation formelle mais ont recours aux caractéristiques de la parole orale dans leurs écrits (ibid., § 22).

Cependant, aucun de ces deux types d'écriture populaire ne reflète la véritable langue parlée orale. Cet état de la langue n'a été écrit qu'à partir du xix^e siècle, lorsque furent publiés des textes accordant délibérément une place importante à l'arabe égyptien. L'écrit-

⁶Pierre Larcher remarque que dans le monde arabe le judéo-arabe est « en général écrit en caractères hébraïques », mais qu'en arabe chrétien, « les textes rédigés en caractères autres qu'arabes sont plutôt rares » (grec, copte, syriaque). (Larcher, 2001, p. 598)

⁷Une forme arabe traditionnelle de poésie strophique orale déclamée dans une langue dialectale.

ture dialectale se caractérise par les thèmes traités dans ces textes et inclut dès le début des thèmes de la vie populaire, d'abord pour parodier le comportement de personnages stéréotypés des classes populaires (paysans, ...) et progressivement au ^{xx}^e siècle pour donner la parole aux ouvriers, aux paysans ou aux révolutionnaires dans des textes aux objectifs progressistes ou contestataires (Doss, 1996, § 33-34).

1.2.2 Le théâtre de la Nahḍa

Avec le début de la période de la *Nahḍa*, «la Renaissance», au ^{xix}^e siècle, on observe qu'un certain nombre d'auteurs se sont mis à utiliser intentionnellement l'arabe égyptien comme support de communication. Le théâtre apparaît comme le domaine naturel de recours à la langue parlée, et le pas est franchi avec les dramaturges suivants :

Muḥammad 'Uṭmān Ġalāl (1829-1898) a traduit et fait jouer des pièces d'auteurs européens, notamment de Molière (Vollers, 1895, p. 3) en arabe égyptien, dont un Tartuffe traduit sous le titre *al-Šeḥ matliḥ*.

Ya'qūb Ṣannū' (1839-1912) a créé sa troupe de théâtre de langue arabe en 1870 pour jouer des comédies qu'il a écrites lui-même, souvent inspirées d'auteurs européens. Les représentations rencontrent un succès immédiat. Il obtient le parrainage du khédive Ismā'īl, qui l'appelle «le Molière de l'Égypte». Il n'hésite pas à avoir recours au dialecte égyptien, ce qui lui attire les critiques de l'éditeur d'un journal italien d'Alexandrie, qui lui reproche de ne pas respecter la grammaire arabe. Il se justifie alors à travers les propos d'un des personnages de sa pièce *Mūlyīr Miṣr wa ma yuqāsīh* (Le Molière égyptien et ce qu'il endure) : "الكوميديّة تشتمل على ما يحصل ويتأتى بين الناس." (La comédie contient ce qui se passe et ce qui survient parmi le peuple.) (Sannū', 2019, p. 15) (Manfred Woidich, 2010, p. 70-71).

On constate donc que dès le début, l'utilisation du dialecte, même pour rendre des dialogues naturels au théâtre, n'échappe pas à la critique. Cette variété de langue est ressentie comme illégitime pour des œuvres littéraires, pour lesquelles l'arabe littéral est la langue attendue. Les auteurs de textes en dialecte se justifient alors en invoquant la nécessité d'utiliser la langue réellement employée par les gens du peuple.

1.2.3 Ouverture du champ littéraire

Maḥmūd Ṭāhir Ḥaqqī (1884-1964) est parmi les premiers à avoir utilisé le dialecte dans les dialogues d'un roman, en l'occurrence *'Aḍrā' Dinšuwāy* 'La Vierge de Dinšuwāy', publié au Caire en 1906, dont la toile de fond est le massacre de paysans du delta par des officiers britanniques. **Muḥammad Ḥusayn Ḥaykal** (1888-1956), fut l'auteur de nombreux romans dont *Zaynab*, *Manāẓir wa-aḥlāq riḥiyya* 'Zaynab, paysages et usages ruraux', paru en 1914. L'auteur y a aussi recours au dialecte pour le discours direct. Dans ce roman, les paysans parlent principalement le dialecte, tandis que la langue des personnages de plus haut rang varie entre 'āmmiyya et fuṣḥā selon l'interlocuteur et le sujet (Diem, 2006, p. 109). **Maḥmūd Taymūr** (1894-1973), écrivain et pionnier de la nouvelle en langue arabe, a aussi écrit en dialecte certains des dialogues de ses premières histoires, notamment dans plusieurs de celles figurant dans le recueil intitulé *al-Šayḥ Ġum'a wa qīṣaṣ 'uḥrā* 'Le Cheikh Gum'a et autres histoires', publié en 1925 (ibid., p. 113).

Des années 1920 aux années 1940, le genre des *muḍakkirāt*, fictions de type autobiographique où un personnage illettré est censé avoir dicté le récit de sa vie à un éditeur, inaugure les récits intégralement écrits en arabe égyptien. Ainsi, **Maḥmūd Bayram al-Tūnisī** (1893-1961) publie en 1925 *is-Sayyid wi mrātu f-Barīs* 'Monsieur et sa femme à Paris' et **Ḥusnī Yūsuf** publie en 1930 *Muzakkirāt fitiwwa* 'Mémoires d'un fier-à-bras'. Cette littérature de genre s'adresse à un public populaire, l'emploi du dialecte la rendant ainsi plus accessible, tout en restant acceptable, tant qu'elle n'empiète pas sur le domaine de la littérature reconnue (Doss, 1996, § 33). Un cortège d'environ trente titres se termine avec *Muḍakkirāt ṭālib bi 'īa* 'Mémoire d'un étudiant boursier', écrit par **Lūwīs 'Awaḍ** (1915-1990), ce récit étant l'autobiographie de l'auteur, et non pas une fiction (ibid., § 45).

Tawfiq al-Ḥakīm (1898-1987), écrivain, dramaturge, essayiste et homme de lettres, fut l'un des géants du roman et de l'écriture théâtrale arabe. Il a utilisé le dialecte pour les dialogues de plusieurs de ses romans, dont *'Awdat al-rūḥ* 'Le Retour de l'esprit', publié en 1933. Il réserva ce choix à ses histoires se déroulant en Égypte. En tant que dramaturge, Tawfiq al-Ḥakīm a écrit plusieurs de ses premières pièces en arabe égyptien, telles que *al-Mar'a al-ḡadīda* (1923), 'La Nouvelle Femme', *Ḥaya taḥaṭṭamat* (1930), 'Une vie s'est effondrée', *Raṣāṣa fi l-qalb* (1931), 'Balle dans le cœur', *az-Zammār* (1932), 'Le flutiste', *Ḥadīṭ ṣuḥufi* (1938), 'Une interview à la presse'. Pour lui, alors que les pièces en arabes littéral étaient surtout destinées à la lecture, celles en dialecte servaient directement pour la représentation (Diem, 2006, p. 120). Avec sa pièce *al-Ṣafqa* (1956), 'Le Contrat', il essaya de créer une langue intermédiaire entre fuṣḥā et 'āmmiyya, tentative qui resta cependant sans lendemain.

Yūsuf 'Idrīs (1927-1991), homme de lettres, nouvelliste et dramaturge, eut recours au dialecte dans les dialogues de certains de ses romans, surtout ceux dont l'intrigue se situe en Égypte, comme *'Arḥaṣ layālī* (1954) 'Les nuits les moins chères'. Il n'a, lui non plus, pas échappé aux reproches. Dans son introduction à *Ġumhūriyyit Faraḥāt* "La République de Faraḥāt", deuxième recueil de 'Idrīs paru en 1957, **Taha Hussein** (1889-1973), le "doyen de la littérature arabe" loue les talents du jeune nouvelliste mais il termine en critiquant son choix de faire parler ses personnages en dialecte : "il est éloquent quand il parle, et s'il fait parler ses personnages, il les fait parler en dialecte, tout comme certains d'entre eux se parlent en réalité lorsqu'ils se rencontrent et discutent les uns avec les autres. Combien nos jeunes écrivains se trompent lorsqu'ils pensent que pour décrire la réalité de la vie, ils doivent faire parler les gens de leurs livres dans la langue des conversations des rues et des clubs."⁸ (Idrīs, 2018, p. 9).

Pourtant, Youssef Idriss a montré sa détermination à revendiquer ce recours au dialecte dès son premier recueil. Le titre *أرخص ليالي* *Arḥaṣ Layālī* aurait dû s'écrire *أرخص ليال* *Arḥaṣ Layālīn* selon les canons de la grammaire arabe classique, mais l'écrivain a insisté pour conserver cette transcription de la prononciation dialectale malgré les critiques de l'époque (al-Naqqach, 1991, p. 151).

⁸ "فهو مفصح إذا تحدث، فإذا أنطق أشخاصه أنطقهم بالعامية كما يتحدث بعضهم إلى بعضني واقع الأمر حين يلتقون ويدرون بينهم. وما أكثر ما يخطئ الشباب من أدبائنا حين يظنون أن تصوير الواقع من الحياة يفرض عليهم أن ينطقوا الناس في الكتب بما تجري به ألسنتهم في أحاديث الشوارع والأندية."

L'ajout du graphe ي final⁹ est donc un choix délibéré visant à marquer l'emploi du nom pluriel indéfini ليالي layālī comme appartenant à la variété de la 'āmmiyya. Les choix d'écriture reviennent chez Youssef Idriss à distancier l'arabe égyptien de l'arabe littéral ne serait-ce que pour revendiquer le droit de transcrire l'arabe dialectal.

À l'instar de Ṭaha Ḥusayn, **Nağīb Maḥfūz** (1911-2006), l'autre géant de la littérature égyptienne, prix Nobel de littérature 1988, s'est aussi opposé à l'utilisation de la 'āmmiyya en littérature. De plus, à l'apogée du panarabisme, notamment après 1952, avec la prise du pouvoir en Égypte par le Mouvement des officiers libres créé par **Gamāl 'Abd al-Nāṣir**, l'unité de la nation arabe est perçue comme devant se faire par le recours à une langue unique, qui ne peut être que la fuṣḥā. L'emploi de la 'āmmiyya est alors ressenti comme une incitation à la division. Les pionniers du dialecte revoient alors leur position : Haykal n'écrit plus ses œuvres qu'en arabe littéral, Taymūr réécrit ses premières nouvelles passant de la 'āmmiyya à la fuṣḥā : ainsi *Abū 'Alī 'āmīl 'artīst* (1934) devient en 1954 *Abū 'Alī 'al-fannān* 'Abū 'Alī l'artiste'. (Diem, 2006, p. 116). Par conséquent, des années 1940 aux années 1990, les romans en dialecte égyptien se font rares. On peut citer de Muṣṭafā Muṣarrafa *Qanṭara al-laḏī Kaḥar* (1965) 'Qantara qui a perdu la foi' (Lagrange, 2013) et de 'Uṭmān Ṣabrī (1896-1986) *Riḥla fī l-Nīl* (1965) 'Voyage sur le Nil' et *Bēt Sirri* (1981) 'Maison close' qui utilisent un mélange d'arabe littéral et d'arabe égyptien.

Par contre, l'emploi de l'arabe égyptien au théâtre devient courant après la révolution égyptienne de 1952, avec Nu'mān 'Aṣūr (1918-1987), 'Alfrīd Farağ (1929-2005), Sa'd al-Dīn Wahba (1925-1997), Raṣād Ruṣḏī et Yūsuf 'Idrīs et plus tard 'Alī Sālim et Nağīb Surūr (Davies, 2005, p. 601).

Finalement dans les années 1990, le recours à la 'āmmiyya chez les auteurs explose et la prose publiée en arabe égyptien comprend l'autobiographie d'Aḥmad Fu'ād Negm al-Fāgūmī, *Tārīx ḥayāt muwāṭin šāyil fī qalbu ... waṭan* (1993) 'Histoire de la vie d'un citoyen portant dans son cœur ... une nation' et les romans suivants : *Laban al-'Aṣṣūr* (1994) 'Le lait de l'oiseau' de Yūsuf al-Qa'īd (Zack, 2001), *Šams il-'Aṣīl* (1998) 'Soleil du soir' de Bahā 'Awwād, *Min Ḥalāwīt il- Rūḥ* (1998) 'Zeste de vie' de Ṣafā 'Abdel Mon'im, *Bānhūf lšīrāsā* (1999) 'Bahnhofstraße' de Sāmiḥ Farağ (Davies, 2005, p. 599).

L'expansion continue au début du xxi^e siècle avec *Tāksī* (2005) 'Taxi' de Ḥalīd al-Ḥamīssī, *Kitāb ma-lūš ism* (2009) 'Un Livre sans nom' de 'Aḥmad al-'Isēlī, ainsi que les autobiographies de Muḥammad Nāṣir 'Alī 'Ūlā 'Awwil (2001) 'Première classe d'école primaire' et de Fathīa al-'Asal *Ḥuḍn il-'Umr* (2004) 'L'étreinte d'une vie' (Manfred Woidich, 2010, p. 82-83). Parmi les expatriés vivants au Caire, Carmine Cartolano a écrit en arabe égyptien *Maṣriyyano : Yawmiyyat muṣawwar 'īṭālī* (2012) 'Journal d'un photographe italien' et *Momo* (2017). De plus en plus de romans sont écrits par des autrices comme par exemple 'Āyzaḥ atgawwiz (2008) 'Je veux me marier!' de Ġāda 'Abdel 'Āl et *Šaklahā sāfarit* (2016) 'Elle a l'air d'avoir voyagé' de Soha Elfeqy. En 2022, une sorte de consécration est atteinte avec la publication du roman *Bil-Ḥīm il-kīnī* 'Sous le sceau kenyan' de Hilāl Šīrīn sous l'égide de l'Autorité générale du livre du ministère égyptien de la Culture.

Un palier est aussi franchi avec la publication en 2016 de l'ouvrage de référence de la littérature arabe *Risālat al-ğufrān* "L'Épître du pardon" de Abū al-'Alā' al-Ma'arrī

⁹En Égypte, le yā final s'écrit presque systématiquement sans ses deux points, comme la 'alif maṣūra ع. Quand ce sera le cas, je respecterai cette graphie.

(973-1057), traduit en arabe égyptien par Nārīmān al-Šāmīlī (al-Kutub Ḥān lil-našr wa al-tawzīʿ, Cairo). Dans son introduction, elle explique qu'en tant que jeune lectrice, elle n'était pas capable de comprendre ce texte classique, ce qui l'a ensuite motivée à le traduire en arabe égyptien, afin de le rendre accessible aux lecteurs égyptiens (Rosenbaum, 2017, p. 59-61).

Par ailleurs, le roman épistolaire *Jawābāt Ḥarājī il-Guṭṭ* (2001) 'Lettres de Haraji le chat' d'Abd ir-Raḥmān il-Abnūdī (1938-2015) est exceptionnel dans son utilisation de l'arabe ṣaʿīdī de Haute Égypte (Davies, 2005, p. 599).

Les revues du xxi^e siècle publiant en arabe égyptien comprennent *Bārī* (à partir d'au moins 2002), l'hebdomadaire satirique *Iḏḥak lil-Dunyā* (à partir de 2005) 'Souris au monde', et le mensuel *Iḥna* (à partir de 2005) 'Nous'.

1.3 Développement de systèmes d'écriture

1.3.1 Adaptation de l'abjad arabe

L'écriture de l'arabe fuṣḥā a d'abord été manuscrite. Des exemples d'impressions en bloc sur papier sont attestés dans l'Égypte fatimide dès le x^e siècle EC, mais l'imprimerie arabe à caractères mobiles est d'abord apparue en Italie au début du xvi^e siècle. Le monde arabo-musulman lui a longtemps préféré la lithographie, plus respectueuse de la calligraphie arabe. L'utilisation ininterrompue de l'imprimerie arabe parmi les Arabes musulmans remonte à 1822, lorsque le premier livre est sorti de la presse d'État de Muḥammad ʿAlī, gouverneur d'Égypte de 1805 à 1848 (Roper, 2014).

Un sentiment d'inadéquation entre le système d'écriture arabe et les nécessités de l'imprimerie est apparu dans les cercles intellectuels arabes. La pluralité de réalisations d'une seule lettre (initiale, médiane, finale, isolée), la gestion des signes diacritiques, notamment, imposaient un grand nombre de types d'imprimerie, rendant l'opération peu pratique.

Dans les années 1930, l'Académie de la langue arabe du Caire a ouvert un concours pour élaborer un procédé idéal de simplification du système d'écriture arabe. En 1944, deux projets sont présentés : celui d'Abd al-ʿAzīz Fahmī inclut le recours à des graphes latins (cf. in 1.3.2 page 29) ; le second, d'Alī l-Ġarīm, conserve les formes originales des consonnes arabes mais ajoute de nouvelles formes de lettres pour les voyelles courtes, la nūn et la madda. Un mois plus tard, l'Académie décide d'abandonner les deux projets, mais néanmoins de les publier, ainsi que les procès-verbaux des discussions (Haralambous, 1998).

En 1959, une sous-commission au sein de l'Académie de la langue arabe du Caire est chargée d'élaborer un projet de simplification typographique de l'arabe, lequel sera approuvé et mis en œuvre en 1960. Ce projet propose une composition de l'arabe avec seulement 72 types (le système traditionnel de base en utilisant environ 470), grâce notamment à la dissociation des ligatures inférieures, réutilisables pour plusieurs lettres dont les réalisations médianes et finales peuvent alors être unifiées (ibid.). Les enjeux existant à l'époque sont désormais quasiment dépassés grâce à la composition informatisée moderne.

Un traitement spécifique pour l'orthographe dialectale arabe est celui présenté par

العدد	حركات قصيرة	أمثلة	حركات ممدودة	أمثلة
١	ا = َ	ma = مَ	ā = َا	mā = مَا
٢	u = ُ	mu = مُ	ū = ُو	mū = مُو
٣	ī = ِ	mi = مِ	ī = ِي	mī = مِي
٤	ä = َ	mä = مَ	ā̇ = َا	mā̇ = مَا
٥	e = ِ	me = مِ	ē = ِي	mē = مِي
٦	o = ُ	mo = مُ	ō = ُو	mō = مُو
٧	ö = ِ	mo = مِ	ō̇ = ِي	mō̇ = مِي
٨	ü = ُ	mü = مُ	ū̇ = ُو	mū̇ = مُو

Fig. 1.1: Préconisations de Ḥalīl 'Asākir pour transcrire les voyelles des dialectes arabes (Asākir, 1955, p. 185)

l'universitaire égyptien Ḥalīl 'Asākir (1906-1993) en 1955 dans *Ṭarīqa li-kitāba nuṣuṣ il-lahğāt il-'arabiyya il-hadīṭa biḥurūf 'arabiyya* 'Méthode d'écriture des dialectes arabes modernes par les lettres arabes', paru dans le numéro 8 du magazine de l'Académie de la langue arabe du Caire. C'est la seule approche de ce type qui ait été approuvée par cette institution. Elle préconise, entre autres, en plus des trois signes de voyelles courtes de la fuṣḥā, l'utilisation de cinq signes diacritiques supplémentaires pour les autres voyelles présentes dans un certain nombre de dialectes du monde arabe, signes à ajouter par-dessus ou par-dessous les lettres arabes standards pour produire leurs formes dialectales (voir tableau 1.1).

Pour les consonnes, Ḥalīl 'Asākir imagine également de nouvelles lettres pour rendre compte de phonèmes absents de la fuṣḥā mais présents dans certains dialectes. En plus de la traditionnelle lettre ġīm avec un point en dessous pour /d͡ʒ/, il suggère sa version avec

prononciation		écriture			dialecte
fuṣḥā	‘āmmiyya	fuṣḥā	phonétique	étymologico-phonétique	
ḍ	g	ج	ج		Le Caire
	ʒ		چ		Syrie, Liban
q	ʔ	ق	أ	ق	Le Caire
	g		ج	ي	Haute-Égypte
θ	s	ث	س	ت	
ð	z	ذ	ز	ت	
ð ^c	z ^c	ظ	ظ		

Fig. 1.2: Préconisations de Ḥalīl ‘Asākir pour transcrire les consonnes des dialectes arabes (Asākir, 1955, p. 189)

deux points pour /g/ et trois points pour /ʒ/. Pour les prononciations alternatives de qāf, il laisse le choix entre une écriture phonétique avec hamza pour le “coup de glotte” /ʔ/¹⁰ ou ġīm avec deux points pour /g/ et une écriture étymologico-phonétique avec qāf surmonté d’une hamza à la place de ses deux points pour /ʔ/ et qāf avec les deux points non plus au dessus mais au dessous. Pour le tā’ prononcé /s/, il propose soit sīn soit tā’ avec trois points horizontaux, pour le dāl prononcé /z/, il propose soit zayn soit dāl avec deux points et quant au zā prononcé /z^c/ il propose de doubler son point (voir tableau 1.2) (Asākir, 1955).

Cette norme ne sera pratiquement pas utilisée en dehors de cercles très restreints (Habash, Diab et Rambow, 2012). On retrouve cependant le qāf surmonté d’une hamza employé dans le dictionnaire de Abd il-Mun‘im Sayyid‘ Abd al-‘Āl intitulé *Mu‘ğam al-Alfāz al-‘Āmmiyya* ‘Dictionnaire de mots dialectaux’ paru en 1971 (Rosenbaum, 2004, p. 286).

¹⁰Consonne occlusive glottale.

De son côté, dans une sorte de postface intitulée «Vers une langue arabe moderne» écrite à la fin de son roman de 1994 *Marāʾi al-qatl* 'Les Pâturages de la tuerie', Fathi Imbābi (né en 1949) a proposé de supprimer les deux points au-dessus du qāf <ق> pour illustrer sa prononciation en arrêt glottal en arabe du Caire par l'utilisation de la graphie <ق> (Imbābi, 1994, p. 414). Il n'est cependant pas allé jusqu'à appliquer cette innovation aux dialogues en arabe égyptien de son roman, et je n'en ai pas trouvé trace ailleurs. Le fait que cette graphie soit dépourvue d'une touche dédiée sur les claviers modernes ne contribue sans doute pas à sa popularisation.

1.3.2 Recours à l'alphabet latin

La seconde moitié du xix^e siècle a été témoin de l'éclosion de l'intérêt d'érudits occidentaux pour l'arabe égyptien, notamment à la suite de l'ouverture du canal de Suez en 1869. La Guerre anglo-égyptienne de 1882 aboutit à la mainmise britannique en Égypte. La présence accrue d'Occidentaux dans le pays encourage l'étude de la langue qui y est parlée selon les critères scientifiques alors en vigueur en Europe. Le besoin de transcrire la prononciation exacte se fait alors particulièrement sentir. Le système d'écriture arabe est alors ressenti comme inadéquat, et le recours à une transcription dans un alphabet latin adapté se présente comme la solution la plus satisfaisante.

La première véritable description détaillée de l'arabe égyptien fut l'œuvre de **Wilhelm Spitta** (1853-1883), orientaliste allemand, directeur de la Bibliothèque vice-royale du Caire, qui publia en 1880 une *Grammatik des arabischen Vulgärdialectes von Aegypten* 'Grammaire du dialecte arabe vulgaire d'Égypte', de 519 pages, dont l'une des principales réalisations est la création d'un système de transcription utilisant un signe par phonème. Il utilise les lettres , <r>, <g>, <d>, <ṛ>, <z>, <s>, <f>, <q>, <k>, <l>, <m>, <n>, <h>, <w> pour les sons similaires en arabe égyptien, un point sous les emphatiques <d>, <t> et <s> (mais en ignorant /zʕ/ ط) et sous le <h> de /ħ/ ح, une brève sous le <h> de /x/ خ, un accent aigu sur le <ś> pour /ʃ/ ش, un point sur le <ḡ> pour /ɣ/ غ et le demi-anneau à gauche pour /ʕ/ ع et à droite pour la hamza /ʔ/. Les voyelles longues sont indiquées par un accent circonflexe, sauf le /i:/ long qui est transcrit par <y>, <j> étant utilisé pour la semi-voyelle correspondante /j/ (Spitta, 1880, p. 1).

Karl Vollers (1857-1909), orientaliste allemand, publia en 1890 un *Lehrbuch der aegypto-arabischen Umgangssprache* 'Manuel de langue familière égypto-arabe' (233 pages), réédité en anglais en 1895 sous le titre *The Modern Egyptian Dialect of Arabic : A Grammar With Exercises, Reading Lessons and Glossaries* (276 pages), traduit par Francis Crawford Burkitt, qui indique dans sa préface : «Le système d'écriture arabe, même entièrement vocalisé, ne donne pas les distinctions les plus fines des voyelles» (Vollers, 1895). Vollers reprend le système de transcription de Spitta (en y incorporant <z> pour /zʕ/ ظ), mais en recourant aux digrammes <ch> pour /x/ خ, <gh> pour /ɣ/ غ et la double apostrophe <'> pour ق prononcé comme la hamza, marquée, elle, par une seule apostrophe (Vollers, 1890, p. 2-3).

Karl Hagenmacher publia en 1892 sa *Grammatik des ägyptisch-arabischen Vulgärdialektes* de 113 pages, qu'il présente dans son introduction comme une alternative pratique et peu chère aux contributions de Spitta et Vollers. Il reprend essentiellement la transcription de Spitta, mais utilise le digramme <kh> pour /x/ خ et le trigramme <sch>

pour /ʃ/ ش.

Friedrich Probst, publia la même année son *Arabischer Sprachführer in ägyptischem Dialect* 'Guide de conversation arabe en dialecte égyptien' de 280 pages, dont le but affiché dans l'introduction est « d'offrir aux voyageurs [...] un moyen de communication avec la population arabophone du nord-est [...] de l'Afrique [...], nul autre que l'égyptien ne peut être utilisé à cette fin. » Il se distingue par l'utilisation des graphes ⟨š⟩ pour /ʃ/ ش, ⟨ç⟩ pour /ʕ/ ع et ⟨ḳ⟩ pour /q/ ق.

En 1895, avec sensiblement le même système de transcription que celui de Spitta, paraît *An Arabic-English vocabulary of the modern and colloquial Arabic of Egypt* (661 pages) et en 1897 sa version *English-Arabic* (554 pages), tous deux compilés par **Socrates Spiro** (né en 1868) Égyptien d'origine syrienne et grecque, alors qu'il travaille au projet de **Daniel Willard Fiske** (1831-1904). Cet universitaire américain, publia en 1897 *An Egyptian alphabet for the Egyptian People* (55 pages). Dans ce livre, il présente un modèle de transcription pour l'écriture du dialecte égyptien dans l'alphabet latin, modèle basé sur la transcription réorganisée et simplifiée de Spitta, auquel il rend un vibrant hommage dans son introduction. Le but est rien moins que de faire adopter cet alphabet en Égypte pour l'enseignement de l'écriture de l'arabe parlé, qui remplacerait l'arabe littéral pour « 'el ḥagât el gedîde, 'les choses nouvelles,' telles que le commerce, la presse et d'autres sujets sans lien avec la religion » (Zack, 2014, p. 4). Malgré la diffusion de tout un matériel de promotion auprès des enseignants égyptiens, cet alphabet ne fut jamais adopté.

Carlo Alfonso Nallino (1872-1938), arabisant italien, publia en 1900 *L'Arabo parlato in Egitto : grammatica, dialoghi e raccolta di circa 6000 vocaboli* 'L'arabe parlé en Égypte : grammaire, dialogues et recueil d'environ 6000 mots' de 386 pages, avec le système de transcription de Spitta, sauf pour le /i:/ long qui est transcrit par ⟨î⟩, ⟨y⟩ étant utilisé pour la semi-voyelle correspondante /j/. La page de couverture indique qu'il s'agit d'une réédition complétée du *Manuale di Arabo volgare* publié en 1888 à Milan par **Rinaldo Slerlich** et **Dib Ḥaddag**, son professeur, dans lequel les digrammes abondent : ⟨hh⟩ pour /ħ/ ح, ⟨kh⟩ pour /x/ خ, ⟨sc⟩ pour /ʃ/ ش, ⟨rh⟩ pour /ɣ/ غ... De plus, /q/ ق est transcrit ⟨g⟩ devant ⟨a⟩, ⟨o⟩, ⟨u⟩ et ⟨gh⟩ devant ⟨i⟩ et ⟨e⟩, à l'italienne.

John Selden Willmore (1856-1931), un des juges de la cour d'appel du Caire, publia en 1901 *The Spoken Arabic of Egypt*, de 394 pages, qu'il décrit dans son introduction comme l'achèvement de l'œuvre de Spitta (Willmore, 1901, p. vii). Willmore utilise le système de transcription de Nallino mais avec les digrammes ⟨kh⟩ pour /x/ خ, ⟨sh⟩ pour /ʃ/ ش et ⟨gh⟩ pour /ɣ/ غ (ibid., p. 1).

Gairdner note que « l'adoption récente par le gouvernement égyptien [du système de transcription] de Willmore pour les noms de lieux dans ses cartes et plans pourrait sembler avoir contribué dans une certaine mesure à normaliser ce système ». La consultation de cartes égyptiennes du début du xx^e siècle renforce cette impression (Zack, 2014, p. 14).

William Henry Temple Gairdner (1873-1928), missionnaire britannique, publie avec **Kurayyim Sallām** en 1916 sa première édition de *Egyptian Colloquial Arabic. A Conversation Grammar* de 208 pages, avec l'intention affichée d'appliquer des méthodes modernes d'enseignement des langues. Pour la transcription, il s'excuse d'ajouter encore un autre système de transcription, inspiré de celui de l'Association phonétique internationale, avec notamment les graphes ⟨□⟩, ⟨ḏ⟩, ⟨ṭ⟩ et ⟨z⟩ pour les emphatiques et ⟨g⟩ pour /ɣ/ غ.

lettres arabes	API	Spitta 1880	Vollers 1890	Probst 1892	Spiro 1895	Fiske 1897	Nallino 1900	Willmore 1901	Gairdner 1916
ب	b	b	b	b	b	b	b	b	b
ت	t	t	t	t	t	t	t	t	t
ج	g	g	g	g	g	g	g	g	g
د	d	d	d	d	d	d	d	d	d
ر	r	r	r	r	r	r	r	r	r
ز	z	z	z	z	z	z	z	z	z
س	s	s	s	s	s	s	s	s	s
ف	f	f	f	f	f	f	f	f	f
ك	k	k	k	k	k	k	k	k	k
ل	l	l	l	l	l	l	l	l	l
م	m	m	m	m	m	m	m	m	m
ن	n	n	n	n	n	n	n	n	n
ه	h	h	h	h	h	h	h	h	h
و	w	w	w	w	w	w	w	w	w
ي	j	j	j	j	j y	j	y	y	j
ص	s ^ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	□
ض	d ^ʕ	d	d	d	d	d	d	d	ḏ
ط	t ^ʕ	t	t	t	t	t	t	t	t
ظ	z ^ʕ		z	d, z	z		z	z	z
ح	ħ	ħ	ħ	ħ	ħ	ħ	ħ	ħ	ħ
خ	x	ħ	ch	ħ	ħ	ħ	ħ	kh	x
ش	ʃ	ʃ	ʃ	ʃ	ʃ	ʃ	ʃ	sh	ʃ
ع	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ
غ	ɣ	ɣ	gh	ɣ	ɣ	ɣ	ɣ	gh	g
ق	q	q	„	ḳ	q	q	q	q	?
ف	ʔ	ʔ	ʔ	—			ʔ	ʔ	ʔ
ا	a:	â	â	â	â	â	â	â	a:
و	u:	û	û	û	û	û	û	û	u:
و	o:	ô	ô	ô	ô	ô	ô	ô	o:
ي	i:	y	î	î	y	y	î	î	i:
ي	e:	ê	ä	ê	ê	ê	ê	ê	e:
ء	a	a	a	a	a	a	a	a	a
ء	u	u	u	u	u	u	u	u	u
ء	o	o	o		o	o		o	o
ء	i	i	i	i	i	i	i	i	i
ء	e	e	e		e	e		e	e

Tab. 1.1: Systèmes de transcription de l'arabe égyptien (fin xix^e, début xx^e siècles)

Beaucoup de ces érudits occidentaux associaient la forte prévalence de l'analphabétisme en Égypte de cette époque à la situation diglossique de la langue arabe. L'abjad arabe étant intrinsèquement associé à la fuṣḥā, considérée comme complexe à maîtriser, son abandon pour le remplacer par un système alphabétique, où les voyelles s'écriraient, semblait un moyen efficace de faire accéder plus de personnes du peuple à la lecture et l'écriture. L'association de l'alphabet latin à la modernité, et de l'abjad arabe au passé était une vision courante à l'époque (Al-Toma, 1961), bien que rejetée par une grande partie des élites intellectuelles arabes.

Une des premières réelles tentatives autochtones de recours à l'alphabet latin fut le projet de l'académicien égyptien 'Abd al-'Azīz Fahmī (1870-1951), projet qu'il présenta en 1944, en réponse au concours ouvert dans les années 1930 par l'Académie de la langue arabe du Caire dans le but de simplifier le système d'écriture arabe. Sa proposition consiste essentiellement en un mélange entre lettres de l'alphabet latin et lettres arabes, dont certaines (les quatre emphatiques) sont écrites en miroir (Fahmī, 1944, p. 153).

De même inspiration, paraît en 1964 le livre *Naḥwa 'abğadiyya ġadīda* 'Vers un nouvel abjad' de 'Uṭmān Ṣabrī, qui propose deux alphabets phonétiques idéaux, l'un qui utilise les lettres de l'alphabet latin, avec des diacritiques pour transcrire les sons sans équivalents courants dans ce système (circonflexe, cédille, ...), l'autre qui recourt à des symboles qui semblent d'inspiration personnelle. Le but est de remplacer l'abjad existant, qui contribuerait à propager l'analphabétisme et à freiner le progrès de l'écriture, des sciences et de la civilisation dans le monde arabe (Al-Sāwī, 2017, p. 15). Aucune suite ne semble avoir été donnée à cette tentative.

La disparité entre les systèmes vocaliques de l'arabe littéral et de ses géolectes a parfois poussé à des solutions radicales, dont celle du poète libanais Ṣa'īd 'Aql (1912-2014), qui a proposé dans les années 1960 d'utiliser l'alphabet latin pour l'arabe libanais, avec expérimentation dans ses recueils de poésie *Yara* (1961) et *Khumasiyyat* (1978) ainsi que dans une édition spéciale du périodique *Lebnaan*. Mais cette expérimentation n'a pas été pérennisée (Habash, Diab et Rambow, 2012).

L'emploi de symboles issus de l'alphabet latin s'est aussi développé alors que s'intensifiait une étude plus approfondie de l'arabe égyptien. Ainsi, divers dictionnaires de cette langue ont recours aux lettres arabes, aux caractères latins ou à une écriture qui associe les systèmes d'écriture, comme *A Dictionary of Egyptian Arabic : Arabic-English* (Badawi et Hinds, 1986), dont chaque entrée commence par une forme lemmatisée en graphes arabes, sans taškīl, suivie d'une transcription en alphabet latin de manière à indiquer la prononciation précise.

Dans la deuxième moitié du xx^e siècle, l'enseignement de l'arabe égyptien à des apprenants allogènes a donné lieu à de nombreuses méthodes, dictionnaires, lexiques ou grammaires. Leurs auteurs ont adapté les systèmes de transcription précédents pour écrire des mots, des phrases, voire pour retranscrire de longs énoncés (par exemple un feuillet radiophonique intitulé *L'intelligence du diable* de Muḥammad Kāmil Ḥasan al-Muḥāmī (1915-1979) et diffusé en 1961 sur les ondes de la Radio du Caire (Jomier et Khouzam, 1973, p. 179-198).

Par ailleurs à la fin du xx^e siècle, le développement des moyens de communication médiée par ordinateur (ou communication virtuelle) a encouragé l'écriture en arabe dialectal, au travers de messages (SMS, emails, ...), interactions sur les réseaux sociaux, forums, tenue de blogs, ... Est apparu alors ce qui a été appelé « franco-arabe » فرانكو عرب, « fran-

١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠ ١١ ١٢ ١٣ ١٤ ١٥ ١٦ ١٧ ١٨ ١٩ ٢٠ ٢١ ٢٢ ٢٣ ٢٤ ٢٥ ٢٦ ٢٧ ٢٨ ٢٩ ٣٠ ٣١ ٣٢ ٣٣ ٣٤ ٣٥ ٣٦ ٣٧ ٣٨ ٣٩ ٤٠ ٤١ ٤٢ ٤٣ ٤٤ ٤٥ ٤٦ ٤٧ ٤٨ ٤٩ ٥٠ ٥١ ٥٢ ٥٣ ٥٤ ٥٥ ٥٦ ٥٧ ٥٨ ٥٩ ٦٠ ٦١ ٦٢ ٦٣ ٦٤ ٦٥ ٦٦ ٦٧ ٦٨ ٦٩ ٧٠ ٧١ ٧٢ ٧٣ ٧٤ ٧٥ ٧٦ ٧٧ ٧٨ ٧٩ ٨٠ ٨١ ٨٢ ٨٣ ٨٤ ٨٥ ٨٦ ٨٧ ٨٨ ٨٩ ٩٠ ٩١ ٩٢ ٩٣ ٩٤ ٩٥ ٩٦ ٩٧ ٩٨ ٩٩ ١٠٠	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠ ١١ ١٢ ١٣ ١٤ ١٥ ١٦ ١٧ ١٨ ١٩ ٢٠ ٢١ ٢٢ ٢٣ ٢٤ ٢٥ ٢٦ ٢٧ ٢٨ ٢٩ ٣٠ ٣١ ٣٢ ٣٣ ٣٤ ٣٥ ٣٦ ٣٧ ٣٨ ٣٩ ٤٠ ٤١ ٤٢ ٤٣ ٤٤ ٤٥ ٤٦ ٤٧ ٤٨ ٤٩ ٥٠ ٥١ ٥٢ ٥٣ ٥٤ ٥٥ ٥٦ ٥٧ ٥٨ ٥٩ ٦٠ ٦١ ٦٢ ٦٣ ٦٤ ٦٥ ٦٦ ٦٧ ٦٨ ٦٩ ٧٠ ٧١ ٧٢ ٧٣ ٧٤ ٧٥ ٧٦ ٧٧ ٧٨ ٧٩ ٨٠ ٨١ ٨٢ ٨٣ ٨٤ ٨٥ ٨٦ ٨٧ ٨٨ ٨٩ ٩٠ ٩١ ٩٢ ٩٣ ٩٤ ٩٥ ٩٦ ٩٧ ٩٨ ٩٩ ١٠٠	١ ٢ ٣ ٤ ٥ ٦ ٧ ٨ ٩ ١٠ ١١ ١٢ ١٣ ١٤ ١٥ ١٦ ١٧ ١٨ ١٩ ٢٠ ٢١ ٢٢ ٢٣ ٢٤ ٢٥ ٢٦ ٢٧ ٢٨ ٢٩ ٣٠ ٣١ ٣٢ ٣٣ ٣٤ ٣٥ ٣٦ ٣٧ ٣٨ ٣٩ ٤٠ ٤١ ٤٢ ٤٣ ٤٤ ٤٥ ٤٦ ٤٧ ٤٨ ٤٩ ٥٠ ٥١ ٥٢ ٥٣ ٥٤ ٥٥ ٥٦ ٥٧ ٥٨ ٥٩ ٦٠ ٦١ ٦٢ ٦٣ ٦٤ ٦٥ ٦٦ ٦٧ ٦٨ ٦٩ ٧٠ ٧١ ٧٢ ٧٣ ٧٤ ٧٥ ٧٦ ٧٧ ٧٨ ٧٩ ٨٠ ٨١ ٨٢ ٨٣ ٨٤ ٨٥ ٨٦ ٨٧ ٨٨ ٨٩ ٩٠ ٩١ ٩٢ ٩٣ ٩٤ ٩٥ ٩٦ ٩٧ ٩٨ ٩٩ ١٠٠
١٠١ ١٠٢ ١٠٣ ١٠٤ ١٠٥ ١٠٦ ١٠٧ ١٠٨ ١٠٩ ١١٠ ١١١ ١١٢ ١١٣ ١١٤ ١١٥ ١١٦ ١١٧ ١١٨ ١١٩ ١٢٠ ١٢١ ١٢٢ ١٢٣ ١٢٤ ١٢٥ ١٢٦ ١٢٧ ١٢٨ ١٢٩ ١٣٠ ١٣١ ١٣٢ ١٣٣ ١٣٤ ١٣٥ ١٣٦ ١٣٧ ١٣٨ ١٣٩ ١٤٠ ١٤١ ١٤٢ ١٤٣ ١٤٤ ١٤٥ ١٤٦ ١٤٧ ١٤٨ ١٤٩ ١٥٠ ١٥١ ١٥٢ ١٥٣ ١٥٤ ١٥٥ ١٥٦ ١٥٧ ١٥٨ ١٥٩ ١٦٠ ١٦١ ١٦٢ ١٦٣ ١٦٤ ١٦٥ ١٦٦ ١٦٧ ١٦٨ ١٦٩ ١٧٠ ١٧١ ١٧٢ ١٧٣ ١٧٤ ١٧٥ ١٧٦ ١٧٧ ١٧٨ ١٧٩ ١٨٠ ١٨١ ١٨٢ ١٨٣ ١٨٤ ١٨٥ ١٨٦ ١٨٧ ١٨٨ ١٨٩ ١٩٠ ١٩١ ١٩٢ ١٩٣ ١٩٤ ١٩٥ ١٩٦ ١٩٧ ١٩٨ ١٩٩ ٢٠٠	٢٠١ ٢٠٢ ٢٠٣ ٢٠٤ ٢٠٥ ٢٠٦ ٢٠٧ ٢٠٨ ٢٠٩ ٢١٠ ٢١١ ٢١٢ ٢١٣ ٢١٤ ٢١٥ ٢١٦ ٢١٧ ٢١٨ ٢١٩ ٢٢٠ ٢٢١ ٢٢٢ ٢٢٣ ٢٢٤ ٢٢٥ ٢٢٦ ٢٢٧ ٢٢٨ ٢٢٩ ٢٣٠ ٢٣١ ٢٣٢ ٢٣٣ ٢٣٤ ٢٣٥ ٢٣٦ ٢٣٧ ٢٣٨ ٢٣٩ ٢٤٠ ٢٤١ ٢٤٢ ٢٤٣ ٢٤٤ ٢٤٥ ٢٤٦ ٢٤٧ ٢٤٨ ٢٤٩ ٢٥٠ ٢٥١ ٢٥٢ ٢٥٣ ٢٥٤ ٢٥٥ ٢٥٦ ٢٥٧ ٢٥٨ ٢٥٩ ٢٦٠ ٢٦١ ٢٦٢ ٢٦٣ ٢٦٤ ٢٦٥ ٢٦٦ ٢٦٧ ٢٦٨ ٢٦٩ ٢٧٠ ٢٧١ ٢٧٢ ٢٧٣ ٢٧٤ ٢٧٥ ٢٧٦ ٢٧٧ ٢٧٨ ٢٧٩ ٢٨٠ ٢٨١ ٢٨٢ ٢٨٣ ٢٨٤ ٢٨٥ ٢٨٦ ٢٨٧ ٢٨٨ ٢٨٩ ٢٩٠ ٢٩١ ٢٩٢ ٢٩٣ ٢٩٤ ٢٩٥ ٢٩٦ ٢٩٧ ٢٩٨ ٢٩٩ ٣٠٠	٣٠١ ٣٠٢ ٣٠٣ ٣٠٤ ٣٠٥ ٣٠٦ ٣٠٧ ٣٠٨ ٣٠٩ ٣١٠ ٣١١ ٣١٢ ٣١٣ ٣١٤ ٣١٥ ٣١٦ ٣١٧ ٣١٨ ٣١٩ ٣٢٠ ٣٢١ ٣٢٢ ٣٢٣ ٣٢٤ ٣٢٥ ٣٢٦ ٣٢٧ ٣٢٨ ٣٢٩ ٣٣٠ ٣٣١ ٣٣٢ ٣٣٣ ٣٣٤ ٣٣٥ ٣٣٦ ٣٣٧ ٣٣٨ ٣٣٩ ٣٤٠ ٣٤١ ٣٤٢ ٣٤٣ ٣٤٤ ٣٤٥ ٣٤٦ ٣٤٧ ٣٤٨ ٣٤٩ ٣٥٠ ٣٥١ ٣٥٢ ٣٥٣ ٣٥٤ ٣٥٥ ٣٥٦ ٣٥٧ ٣٥٨ ٣٥٩ ٣٦٠ ٣٦١ ٣٦٢ ٣٦٣ ٣٦٤ ٣٦٥ ٣٦٦ ٣٦٧ ٣٦٨ ٣٦٩ ٣٧٠ ٣٧١ ٣٧٢ ٣٧٣ ٣٧٤ ٣٧٥ ٣٧٦ ٣٧٧ ٣٧٨ ٣٧٩ ٣٨٠ ٣٨١ ٣٨٢ ٣٨٣ ٣٨٤ ٣٨٥ ٣٨٦ ٣٨٧ ٣٨٨ ٣٨٩ ٣٩٠ ٣٩١ ٣٩٢ ٣٩٣ ٣٩٤ ٣٩٥ ٣٩٦ ٣٩٧ ٣٩٨ ٣٩٩ ٤٠٠

Fig. 1.3: Préconisations de ‘Abd al-‘Azīz Fahmī (Fahmī, 1944, p. 153)

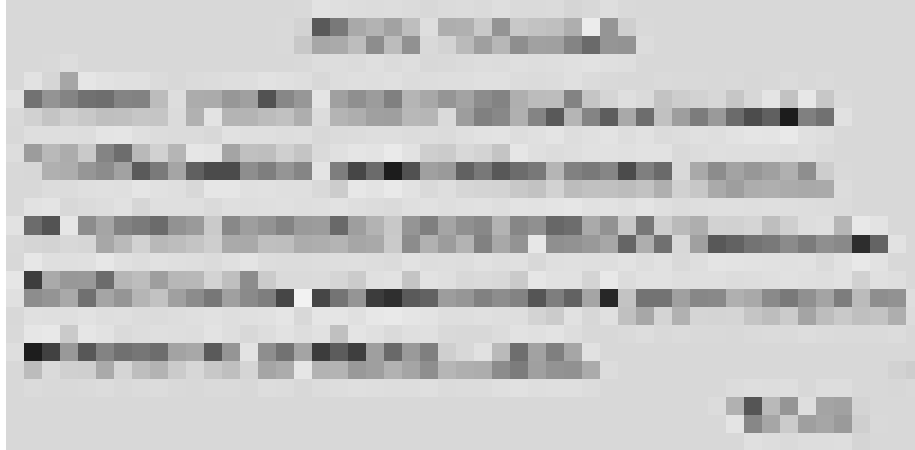


Fig. 1.4: Sūratu l-fātiḥa écrite selon le premier système de ‘Uṭmān Ṣabrī (Al-Sāwī, 2017, p. 16)

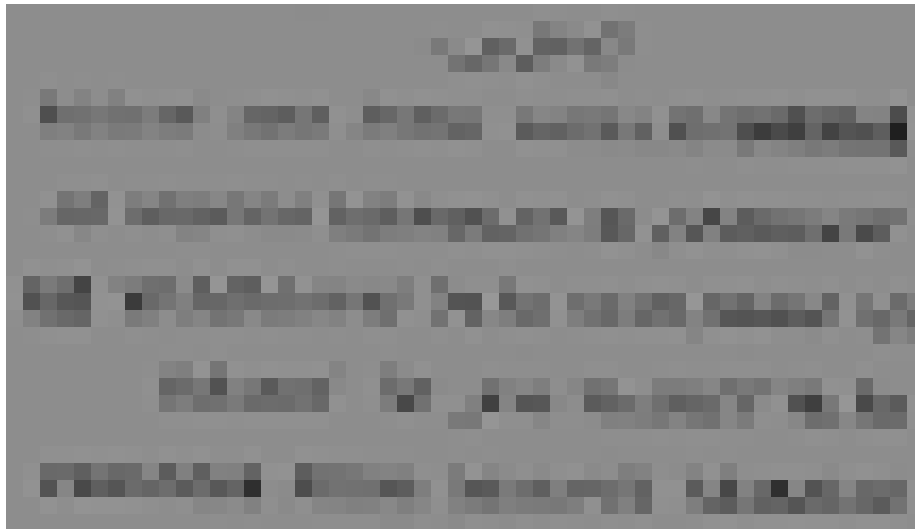


Fig. 1.5: Āyatu l-kursī écrite selon le deuxième système de ‘Uṭmān Ṣabrī (Al-Sāwī, 2017, p. 19)

q̣	=	ق	z	=	ز	q̣	=	ع
q	=	ك	s	=	س	b	=	ب
l	=	ل	c	=	ش	t	=	ت
m	=	م	ṣ	=	ص	ṿ	=	ث
n	=	ن	ç	=	ض	j	=	ج
h	=	ه	ṭ	=	ط	x	=	خ
w	=	و	ẓ	=	ظ	k	=	ك
y	=	ي	y	=	ع	d	=	د
p	=	پ	g	=	غ	ẓ	=	ز
v	=	ف	f	=	ف	r	=	ر
ḥ	=	ح						

Fig. 1.6: Système d'écriture proposé par Şa'īd 'Aql (Al-Sāwī, 2017, p. 21)

co»¹¹ ou encore «arabizi» (mot valise issu de la fusion de 'arabī et 'inglīzī = anglais, ou «easy»), c'est-à-dire l'utilisation de lettres latines pour rendre les phonèmes arabes avec recours aux chiffres pour les cas où l'alphabet latin n'offre pas de correspondance¹². Ce système, qui a émergé spontanément, fait face aux mêmes défis que ceux déjà décrits. Alors que la transcription de certains phonèmes, communs entre les systèmes phonologiques arabes et anglais ou français, semble faire consensus, les utilisateurs ont du mal à trouver un terrain d'entente en ce qui concerne ceux plus spécifiques à l'arabe (Björnsson, 2010, p. 78).

Comme on peut le voir dans les tableaux 1.1 et 1.2, les transcriptions utilisées ont en commun de toutes avoir recours aux mêmes graphes de l'alphabet latin pour une quinzaine de consonnes dont la prononciation est très peu différente entre l'anglais ou le français d'une part, et l'arabe d'autre part. Les difficultés commencent avec les phonèmes consonantiques emphatiques, qui donnent lieu à plusieurs types de graphes, généralement ceux des consonnes non emphatiques, munis de signes diacritiques ou de queues allongées. À noter que l'arabizi relevé par Björnsson utilise un graphe simple pour chaque paire de phonèmes emphatiques, non emphatiques, ce qui semble impliquer qu'il y a peu de cas d'ambiguïté avérés.

¹¹Ce phénomène est dénoncé par 'Ammār al- Ġazzār dans un livre de 2016 intitulé *Frānku 'arab luġa al-šabāb al-hadītha* - فرانكو عرب لغة الشباب الحديثة 'Franco-arabe, langue moderne des jeunes' (Dār al-kutub al-miṣriyya)

¹²Il existe un phénomène équivalent pour le grec avec le greeklish, mot-valise formé sur les mots anglais Greek (grec) et English (anglais).

lettres arabes	API	Mitchell 1956	Badawi et al. 1986	Louis 2008	Abdel- Massih 2009	al- Tonsi 2010	Jomier 1973	Boutros 1993	Girod 2016	Björnsson (arabizi) 2010
ب	b	b	b	b	b	b	b	b	b	b
ت	t	t	t	t	t	t	t	t	t	t
ج	g	g	g	g	g	g	g	g	g	g
د	ʒ	d	j	d	ž	d	d	j	ǰ	d
ر	r	r	r	r	r	r	r	r	r	r
ز	z	z	z	z	z	z	z	z	z	z
س	s	s	s	s	s	s	s	s	s	s
ف	f	f	f	f	f	f	f	f	f	f
ك	k	k	k	k	k	k	k	k	k	k
ل	l	l	l	l	l	l	l	l	l	l
م	m	m	m	m	m	m	m	m	m	m
ن	n	n	n	n	n	n	n	n	n	n
ه	h	h	h	h	h	h	h	h	h	h
و	w	w	w	w	w	w	w	w	w	w
ي	j	y	y	y	y	y	y	y	y	y
ص	s ^ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	ʕ	s
ض	d ^ʕ	ḏ	ḏ	ḏ	ḏ	ḏ	ḏ	ḏ	ḏ	d
ط	t ^ʕ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	ṭ	t
ظ	z ^ʕ	ẓ	ẓ	ẓ	ẓ	ẓ	ẓ	ẓ	ẓ	z
ح	ḥ	ḥ	ḥ	ḥ	ḥ	ḥ	ḥ	ḥ	ḥ	7, (h)
خ	x	x	x	kh	x	x	ḫ	x	ḫ	5, 7', kh
ش	ʃ	ʃ	ʃ	sh	š	ʃ	š	š	š	sh
ع	ʕ	ε	ε	ε	ʕ	ʕ	ε	ε	ʕ	3
غ	ɣ	ɣ	ɣ	gh	ɣ	ɣ	ġ	ġ	ġ	3', gh
ق	q	q	q	q	q	q	q	q	q	k, q
ف	ʔ		-l	ā	ʔ		q	q	q	2, q
ء	ʔ	ʕ	ʕ	ā ē ō ū	ʔ	ʔ	,	,	,	2
ا	a:	áa	aa	a	aa	aa	â	ā	ā	a, aa
و	u:	úu	uu	uu	uu	uu	û	ū	ū	o, oo
و	o:	óó	oo	oo	oo	oo	ô	ō	ō	oo, ou
ي	i:	íi	ii	ii	ii	ii	î	ī	ī	ee, e
ي	e:	ée	ee	ee	ee	ee	ê	ē	ē	e, ee
ء	a	a	a	a	a	a	a	a	a	a, -
ء	u	u	u	u	u	u	u	u	u	u, -
ء	o	o	o	o	o		o	o	o	
ء	i	i	i	i	i	i	i	i	i	i, -
ء	e	e	e	e	e		e	e	e	

Tab. 1.2: Systèmes de transcription de l'arabe égyptien (fin xx^e, début xxi^e siècles)

La consonne fricative pharyngale sourde /ħ/ appelle l'utilisation de variations du graphe ⟨h⟩, avec queue, diacritique, etc. L'arabizi se voit obligé d'utiliser le chiffre 7 (de par sa ressemblance avec la lettre ح) ou simplement le graphe ⟨h⟩, avec risque d'ambiguïté avec la glottale /h/. Des digrammes sont parfois la solution adoptée, comme ⟨sh⟩ pour la consonne fricative palato-alvéolaire sourde /ʃ/, ⟨kh⟩ pour la vélaire sourde /x/ et ⟨gh⟩ pour sa jumelle voisée /ɣ/. Le risque, bien sûr, est l'ambiguïté qui existe avec une suite de deux phonèmes, comme dans /ʔashal/ (plus facile), dont le digramme ⟨sh⟩ transcrit deux phonèmes et doit se prononcer /sh/ et non /ʃ/.

Le graphe ق (qāf) n'est pas univoque en arabe du Caire, car sa prononciation /q/ de consonne occlusive uvulaire sourde qui subsiste en arabe littéral, est généralement remplacée par un "coup de glotte", habituellement notée par une hamza ʾ. Les transcriptions gardent en général le graphe ⟨q⟩ pour la prononciation traditionnelle, et adoptent le même graphe modifié (en exposant, ou réduit, ...), quand ils souhaitent distinguer les deux valeurs.

Pour les voyelles, l'enjeu principal est de différencier entre voyelles longues et voyelles courtes. Le choix de doubler les graphes correspondants est une solution possible avec les claviers les plus simples, d'où son utilisation en arabizi, au moins pour les phonèmes /e:/ et /i:/ (doublement du «e» probablement influencé par l'orthographe anglaise, comme dans «see»). Le phénomène d'anaptyxe se réalise par l'insertion d'une voyelle euphonique rendue nécessaire pour éviter une suite de trois consonnes, impossible en arabe égyptien. Il est rarement retranscrit, car probablement non ressenti comme essentiel, même si cette voyelle peut porter l'accent.

Plusieurs manières d'écrire l'arabe égyptien avec des graphes issus de l'alphabet latin ont donc été imaginées, mais parmi les obstacles principaux à leur adoption figurent le manque de consensus dans le choix des graphes, la complexité de certains d'entre eux et surtout la plus grande familiarité de la vaste majorité des Égyptiens avec les graphes arabes.

1.4 Conclusion

L'arabe égyptien apparaît donc comme une variété d'arabe, parmi de nombreuses autres. Pensé comme un dialecte oral, il s'écrit pourtant depuis très longtemps, initialement de façon spontanée mais désormais de manière de plus en plus revendiquée. Cette prétention au droit d'être écrit implique la nécessité de faire des choix, notamment celui des graphes auxquels recourir. Le défi consistant à transcrire une variété de langue spécifique par un système d'écriture arabe non optimisé pour sa phonologie a encouragé l'apparition d'initiatives innovantes, que ce soit avec des graphes arabes ou avec des symboles issus de l'alphabet latin. Le fait est qu'aucune solution à la fois univoque et pratique n'a été trouvée. Les graphes issus de l'alphabet latin peuvent apparaître comme pertinents pour leur capacité à transcrire précisément la phonologie de l'arabe égyptien, ou à l'inverse pour leur disponibilité et leur caractère universel sur les moyens de communication médiée par ordinateur. Cependant, pour de nombreux Égyptiens, le fait d'aborder l'écriture par l'apprentissage des graphes arabes entraîne une plus grande aisance face à ceux-ci par rapport aux graphes latins. L'arabe égyptien est par conséquent très souvent et probablement systématiquement publié dans une écriture par graphes arabes. Mais en réalité,

tout système d'écriture est un compromis entre les deux idéaux d'univocité et praticité, y compris le système sacralisé par la fuṣḥā.

Chapitre 2

Polyvalences de certains graphes arabes

L'écriture de l'arabe égyptien par des graphes arabes implique le recours à un système optimisé pour écrire et lire l'arabe fuṣṣḥā. Il existe en effet une correspondance presque systématiquement univoque entre graphèmes et phonèmes de la fuṣṣḥā. C'est le cas pour les consonnes du tableau 2.1.

lettre	ب	ت	ث	ج	ح	خ	ذ	ر	ز	س	ش	
API	b	t	θ	dʒ	ħ	x	ð	r	z	s	ʃ	
lettre	ص	ض	ط	ظ	ع	غ	ق	ك	ل	م	ن	ه
API	sˤ	dˤ	tˤ	zˤ	ʕ	ɣ	q	k	l	m	n	h

Tab. 2.1: Lettres et phonèmes arabes en relation isomorphe

Toutefois, quelques exceptions à la règle de l'univocité subsistent. Nous étudierons ci-dessous la 'alif et ses différents emplois, la hamza, la tā' marbūṭa, les wāw et yā', le lām, la šadda et les phénomènes d'assimilation transcrits et non transcrits. Nous souhaitons ainsi décrire les principaux écarts à la règle de l'univocité dans l'emploi des graphes arabes pour la fuṣṣḥā, écarts qui auront des répercussions sur l'écriture de l'arabe égyptien.

2.1 'alif

La première lettre de l'abjad arabe, écrite en position isolée ا, a hérité de deux valeurs phonologiques dans son parcours depuis ses origines.

- la consonne "coup de glotte" /ʔ/, sa valeur d'origine
- la voyelle longue /a:/

Ceci contribue à expliquer la multiplicité de ses emplois contemporains.

2.1.1 'alif et transcription de la voyelle longue /a:/

Pour transcrire la voyelle longue /a:/, la 'alif ا est le graphe prépondérant en arabe fuṣḥā. Dans un texte vocalisé, il est précédé par une consonne portant le signe diacritique fatha (qui transcrit la voyelle courte /a/), et est considéré comme sa prolongation¹. La 'alif transcription de la voyelle longue /a:/ est notamment utilisé dans les schèmes des verbes de formes iii فاعل et vi تفاعل, dans les maṣdārs des verbes de formes iii مفاعلة/فعال, iv إفعال, vi إفاعل, vii انفعال, viii افتعال, ix افعال et x استفعال, dans les participes de forme i فاعل, iii فاعل et vi متفاعل, dans les schèmes nominaux singuliers فعال, فعّال, فاعول, مفعّال, فعّالان, les pluriels أفعال, مفعّال, مفاعيل, فواعل, فواعيل, فعلاء, les terminaisons du duel des verbes لـ/ان, et des noms au nominatif لـ/ان, le pluriel féminin لـ/ات, la 3^e personne de l'accompli des verbes concaves نام, etc.

La voyelle longue /a:/ 'n'est pas transcrite par une 'alif dans quelques de mots, certains étant cependant très courants : إله /ʔila:h/ 'dieu' et ses dérivés, dont الله /alla:h/ 'Dieu', هذا raḥma:n/ 'miséricordieux', les démonstratifs de proximité singuliers masculin هَذَا /ha:ða:/, féminin هَذِهِ /ha:ði:/:, pluriel هَؤُلَاءِ /ha:ʔula:ʔi/, le démonstratif d'éloignement masculin singulier ذَلِكَ /ða:lika/, les conjonctions لَكِنَّ /la:kin/ et لَكِنَّ /la:kinna/, etc. Pour les textes non vocalisés, aucun graphe n'apparaît. Un signe diacritique spécifique est parfois utilisé pour les textes vocalisés (essentiellement dans le Coran); il s'agit de la 'alif suscrite (en arabe : 'alif ḥanjariyya أَلِفْ حَنْجَرِيَّةٌ 'alif poignard'), qui apparaît dans les exemples ci-dessus, et qui consiste en un court trait vertical au-dessus de la consonne précédant /a:/. Dans d'autres vocalisations de textes, on trouve aussi le signe diacritique fatha (qui transcrit habituellement la voyelle courte /a/), surmonté ou nom de la 'alif suscrite.

2.1.1.1 'alif maqṣūra

À la fin d'un mot, la voyelle longue /a:/ est très souvent transcrite par un yā sans point ي, appelé 'alif maqṣūra 'alif courte' comme dans عَلَى /ʕala:/ 'sur'. À l'instar de la 'alif traditionnelle, elle est précédée par une consonne portant dans les textes vocalisés le signe diacritique fatha, dont elle est la prolongation أَلْفَتَى /alfata:/ 'le jeune homme'.

La 'alif maqṣūra est caractéristique du verbe «défectueux» (ناقص) actif, si sa troisième radicale est étymologiquement yā ي. C'est le cas en particulier à la 3^e personne du masculin singulier de l'accompli (māḍī) pour les verbes conjugués sur le modèle بَنَى /bana:/ 'il construisit' et aux personnes du singulier et à la 1^{re} du pluriel à l'inaccompli (muḍāri) pour ceux conjugués sur les modèles suivants pour les formes i يَبْقَى /jabqa:/ 'il reste', v يَتَغَدَّى /jataḡaḍḍa:/ 'il déjeune' et vi يَتَلَا شَىْ /jatala:ʃa:/ 'il disparaît', etc.².

¹ Tout comme la voyelle longue /u:/ est marquée par un wāw و considéré comme la prolongation de la ḍamma /u/ et la voyelle longue /i:/ est marquée par un yā ي considéré comme la prolongation de la kasra /i/

² La finale /aja/ de l'ancien arabe /nahaja/ est devenu /e:/ > /nahe:/ en arabe hijazi (arabe coranique) puis a fusionné avec /a:/ en arabe classique pour donner /naha:/ 'interdire'. Avant l'introduction du système de i'jām, c'est-à-dire les points utilisés pour distinguer les différentes consonnes, les premiers scribes du Coran utilisaient ي pour représenter /j/, /i:/ et /e:/. Au moment de l'apparition du i'jām, /e:/ avait fusionné avec /a:/ et les scribes

Si la troisième radicale des verbes «défectueux» actifs de la première forme est étymologiquement wāw و, on trouve une 'alif traditionnelle ʾ, aussi appelée 'alif ṭawīla 'ʾalif longue', à la 3e personne du masculin singulier à l'accompli : شَكَا /ʃaka:/ 'il se plaignit'. Mais pour tous ces verbes, au passif, on trouve un yā' (avec deux points) à l'accompli : بُنِيَ /bunija/ 'il fut construit', شُكِيَ /ʃukija/ 'il fut plaint' et la 'alif maqṣūra (sans point) à l'inaccompli يَبْنِي /jubna:/ 'il est construit', يَشْكِي /juʃka:/ 'il est plaint'.

Cette 'alif finale, qu'elle soit transcrite par la 'alif maqṣūra ى ou la 'alif ṭawīla ʾ, est appelée 'alif layyina لينة 'alif flexible'. Quand celle-ci est la quatrième lettre, ou plus, dans un mot, la règle veut qu'elle s'écrive ى, comme pour اِشْتَكَى /iʃtaka:/ 'il se plaignit', مِصْطَفَى /musṭafa:/ 'choisi', etc., sauf si elle est précédée de yā' (pour éviter la suite de deux lettres semblables) comme dans دُنْيَا /dunja:/ 'monde', يَحْيَا /jahja:/ 'il vit', l'exception de l'exception étant le prénom يَحْيَى /jahja:/, afin de marquer la différence avec le verbe.

On trouve la 'alif maqṣūra à la fin de noms d'origine étrangère et ancienne : مُوسِيقَى /mu:si:qa:/ 'musique', مُوسَى /mu:sa:/ 'Moïse', عِيسَى /ʕi:sa:/ 'Jésus', etc.

Si, à un nom terminé par une 'alif maqṣūra, on ajoute un suffixe, elle se transforme en 'alif traditionnelle : pronom clitique مُوسِيقَاهُمْ /mu:si:qa:hum/ 'leur musique', tā' marbūṭa اَلْفَتَاةُ /alfata:tu/ 'la jeune femme'. Cependant, l'ajout d'un pronom clitique aux prépositions عَلَى /ʕala:/ 'sur' et إِلَى /ʔila:/ 'vers' transforme la 'alif maqṣūra en yā' عَلَيْهِ /ʕalajhi/ 'sur lui' et إِلَيْهَا /ʔilajha:/ 'vers elle', ce qui rappelle le lien étymologique entre les deux.

2.1.2 'alif orthographique

La 'alif a aussi des emplois autres que la transcription de la voyelle longue /a:/, où elle n'a essentiellement qu'une valeur orthographique, sans réalisation phonologique :

- al-'alif al-fāriqa اَلْف الفارقة 'alif séparatrice' placée après le wāw و final des formes plurielles du verbe à la 3e personne masculin de l'accompli (māḍī) فَعَلُوا /faʕalu:/ 'ils ont fait' de l'impératif ('amr) اِفْعَلُوا /ifʕalu:/ du subjonctif (maṣbūb) et de l'apocopé (majzūm) يَفْعَلُوا /jaʕalu:/ ainsi que la 2e personne masculin du subjonctif et de l'apocopé تَفْعَلُوا /taʕalu:/. Cette 'alif dite orthographique ne se prononce pas et disparaît à l'écrit si un pronom clitique est suffixé au verbe : فَعَلُوهُ /faʕalu:hu/ 'ils l'ont fait'.
- 'alif al-tanwīn اَلْف التنوين qui porte la fathātān ً, fathā redoublée marquant l'indétermination du mot singulier et pluriel non diptote au cas accusatif اَيْضًا /ʔajdaʕan/ 'aussi'. C'est la quasi-unique marque de déclinaison des singuliers dans un texte non vocalisé³. Cette 'alif support ne s'écrit pas après les noms terminés par une tā' marbūṭa

ont donc gardé ى sans points pour le transcrire, et ي avec deux points pour /j/ et /i:/.

³Les cas (nominatifs, accusatifs, génitifs) des six noms أَب /ʔab/ 'père', أَخ /ʔax/ 'frère', حَم /ham/ 'beau-père', فَم /fam/ 'bouche', هَن /han/ 'truc', دُو /ðu:/ 'ayant, doué de' sont respectivement marqués au singulier déterminé par annexion par les voyelles longues wāw و, 'alif ʾ et yā' ي. Au duel et au pluriel externe masculin, les cas sont respectivement différenciés entre nominatif نُون /nun/ d'une part et accusatif/génitif d'autre part يَنْ /yan/.

لُغَةً /luḡatan/ ‘langage’, une hamza simple ماءً /maʔan/ ‘eau’, une hamza sur ‘alif
 خَطَأً /xaṭaʔan/ ‘erreur’, une ‘alif ṭawīla عَصَا /ʕaṣa/ ‘bâton’ ou une ‘alif maqṣūra
 فَتًى /fatan/ ‘un jeune homme’⁴.

- al-’alif al-mamdūda الممدودة la ‘alif allongée’, support de madda آ ‘allongement’, soit un signe placé au dessus de la ‘alif pour indiquer qu’il faut la prononcer comme un “coup de glotte” ayant pour voyelle /a:/ long : آب /ʔa:b/ ‘août’, قرآن /qurʔa:n/ ‘Coran’. Il permet d’éviter d’écrire deux ‘alifs à la suite, soit dans le cas où la première porte une hamza et est suivie d’une ‘alif voyelle longue تَأْكَلْ > تَأْكَلْ /taʔa:kala/ ‘il se corrodait’ ou dans le cas où chacune porte une hamza, la première portant une fatha et la seconde un sukūn أَكُلْ > أَكُلْ /ʔaʔkulu/ > /ʔa:kulu/ ‘je mange’.
- al-’alif al-mahmūza المهمزة la ‘alif hamzée, support de la hamzat al-qaṭ⁶ : c’est le seul support possible de la hamza en début de mot⁷, celle-ci étant placée au dessus de la ‘alif si elle porte la voyelle fatha ا /a/ أرض /ʔard/ ‘terre’ ou ḍamma /u/ أم /ʔumm/ ‘mère’ et au dessous de la ‘alif si elle porte la voyelle kasra /i/ إلى /ʔila:/ ‘vers’. En milieu de mot, la ‘alif est support de hamza seulement si la consonne qui la précède porte la voyelle fatha et si la hamza porte la fatha سَأَلَ /saʔala/ ‘il demanda’ ou le sukūn رَأْسُ /raʔs/ ‘tête’, ou alors si la consonne qui le précède porte le sukūn et la ‘alif porte la fatha مَسْأَلَةٌ /masʔala/ ‘question’. En fin de mot, la ‘alif est support de hamza seulement si la consonne qui la précède porte la voyelle fatha بَدَأَ /badaʔa/ ‘il commença’. Dans les autres cas de vocalisation, la hamza se pose sur la lettre wāw و, le yā ي sans point ى ou seule sur la ligne ء selon les règles complexes exposées page 40 dans la partie consacrée à la hamza.
- la ‘alif support de la hamzat al-waṣl همزة الوصل, appelée aussi hamza instable : elle n’apparaît qu’en position initiale du mot et la hamzat al-waṣl est essentiellement transcrite, dans les textes ornements, par un diacritique en forme de ص (pour وصل = waṣl, lien) : أ. Cette ‘alif indique une liaison avec la voyelle finale du mot précédent. Elle se place avant une suite de deux consonnes, qui, si elle se trouve en début de phrase, après une pause, ne peut se prononcer en prosodie arabe. Dans ce cas, la hamzat al-waṣl transcrit un “coup de glotte”, similaire à celui transcrit par

⁴ le nūn final ن disparaissant devant une annexion ou un pronom clitique.

⁵ À noter que lorsque عَصَا /ʕaṣa/ ou فَتًى /fatan/ sont déterminés, la ‘alif et la ‘alif maqṣūra se prononcent comme la voyelle longue /a:/ العَصَا /al-ʕaṣa:/ ou الفَتَى /al-fata:/. Ces mots n’ont d’ailleurs que ces deux prononciations au singulier, quels que soient les cas de déclinaison.

⁶ Signe posé sur une consonne non suivie d’une voyelle.

⁷ La hamzat al-qaṭ, de «coupure», car elle transcrit un “coup de glotte”, est opposée à la hamzat al-waṣl, de “lien”, qui indique la liaison faite avec la voyelle précédente (cf. item 2.1.2).

Les noms étrangers commençant par une voyelle sont régulièrement rendus en arabe écrit par un mot commençant par une hamza prolongée par أوسلو /ʕu:slu:/ ‘Oslo’, أوروبا /ʕu:ru:bba/ ‘Europe’ ou إيران ي /ʕi:ra:n/ ‘Iran’. Si la voyelle est /a/, on trouve une hamza + fatha أمريكا /ʕamri:ka:/ ‘Amérique’ voire une madda آسيا /ʕa:sja:/ ‘Asie’.

la hamzat al-qaṭʿ, sans que le diacritique apparaisse. Pour les textes vocalisés, on place juste la voyelle courte, généralement kasra sous la ʾalif. Son emploi le plus courant est comme première lettre de l'article défini ال (cf. infra 2.5.2 page 43, on la retrouve dans les relatifs de type الذي). On la trouve aussi en initiale des impératifs de première forme اِفْعَلْ et des formes vii اِنْفَعِلْ, viii اِفْتَعِلْ, ix اِفْعَلْ et x اِسْتَفْعِلْ, de l'accompli des formes vii اِنْفَعَلْ, viii اِفْتَعَلْ, ix اِفْعَلْ et x اِسْتَفْعَلْ, pour certains noms, parmi lesquels ابن /ʔi)bn/ 'fils' et اسم /ʔi)sm/ 'nom' ainsi que des noms d'origine étrangère comme استوديو /ʔi)stu:dju:/ 'studio' qui commencent par une suite de deux consonnes. Le rôle prothétique de la hamzat al-waṣl apparaît clairement dans ces cas.

La différenciation entre la hamzat al-qaṭʿ et la hamzat al-waṣl est souvent objet de confusion. C'est notamment le cas entre le maṣdār des verbes de forme iv اِفْعَال qui commence par une hamzat al-qaṭʿ, et les maṣdārs des verbes de forme vii اِنْفَعَال, viii اِفْتَعَال, ix اِفْعَال, et x اِسْتَفْعَال qui commencent par une hamzat al-waṣl. Ainsi Wright note :

In more modern Arabic, the elision of the élif conjunctionis [...] is neglected, especially after the article as اَلْاِفْتِدَارُ [...]; but the grammarians brand this as ʿhurūḡ ʿan kalām al-ʿarab wa laḥn fāḥiṣʿ.

En arabe plus moderne, l'élision de l'élif conjunctionis [ʾalif support de la hamzat al-waṣl] est négligée, surtout après l'article comme pour اَلْاِفْتِدَارُ [/ʔalʔiqtida:ru/ le pouvoir]; mais les grammairiens qualifient cela « d'écart par rapport à la langue des Arabes et d'erreur obscène ». (Wright, 1967, p. 20)

Il s'avère d'ailleurs que les confusions entre ا et او ou ا et ا sont parmi les erreurs les plus courantes en arabe contemporain (Attia et al., 2015, p. 13).

Conclusion : la ʾalif a donc de multiples rôles dans l'écriture de l'arabe fuṣḥā, et cette plurivocité rend son interprétation en lecture essentielle pour comprendre un texte. Le scripteur doit donc souvent choisir entre l'ajout de signes diacritiques (hamza, fathatān, madda, voyelle courte, ...) et l'urgence de l'écriture, voire la nécessité de ne pas surcharger le texte final.

2.2 hamza

Le "coup de glotte" ou consonne occlusive glottale /ʔ/ est transcrit en arabe soit par le diacritique madda (cf. supra 2.1.2, page 39), soit par le signe nommé hamza ء, qui, bien qu'il corresponde à une des 28 consonnes pouvant occuper une des trois places des racines trilitères, n'est pas considéré comme une lettre. Le graphe ʾalif, qui à l'origine transcrivait ce phonème, ayant acquis la valeur de la voyelle longue /a:/, un rôle de support de diacritiques, voire une fonction orthographique, la hamza, un graphe nouveau, fut créée et essentiellement utilisée en position diacritique.

Son arrivée tardive dans le système d'écriture arabe a eu pour conséquences une transcription différenciée selon les contextes. Comme nous l'avons vu avec la ʾalif support de

la hamzat al-qaṭ‘, celle-ci a pour support une ‘alif systématiquement en début de mot. En milieu de mot, le support dépend du contexte vocalique. En gros, si la hamza est précédée de ou porte la voyelle kasra /i/, elle sera posée sur un yā’ sans point : بِئْر /biʔr/ ‘puits’, sinon, si la hamza est précédée de ou porte la voyelle ḍamma /u/, elle sera posée sur un wāw : بُؤْس /buʔs/ ‘malheur’, sinon, si la hamza est précédée d’une ‘alif, elle sera posée sans support sur la ligne : سَاءَلَ /sa:ʔala/ ‘il interrogea’, sinon elle sera posée sur une ‘alif رَأْسَ /raʔs/ ‘tête’. Même influence du contexte vocalique en fin de mot : derrière une kasra, la hamza sera posée sur un yā’ sans point هَادِيْ /ha:diʔ/ ‘calme’, derrière une ḍamma, sur un wāw جَرُوْ /ʒaruʔa/ ‘il osa’, derrière une faṭḥa, sur une ‘alif بَدَأَ /badaʔa/ ‘il commença’, derrière un sukūn, sans support sur la ligne جَزْءَ /ʒuzʔ/ ‘morceau’, ce qui inclut les voyelles longues, considérées comme portant un sukūn en prolongation de la voyelle courte précédente : بَرِيْءَ /bari:ʔ/ ‘innocent’, سُوءَ /su:ʔ/ ‘mal’, مَاءَ /ma:ʔ/ ‘eau’.

La notation atypique et complexe de la hamza déroge à la règle générale d’univocité dans l’usage de l’arabe fuṣṣḥā, mais elle offre une des rares occasions où les voyelles courtes sont, ne serait-ce que partiellement, matérialisées dans l’écriture non vocalisée.

Cette complexité a cependant donné lieu à des désaccords et des exceptions à la règle (Parkinson, 1990). Ainsi, l’Académie de la langue arabe du Caire, souhaitant éviter la suite de deux و pourtant attestée en arabe classique, prescrit سُئُونِ plutôt que سُؤُونِ /ʃuʔu:n/ ‘sujets, questions’ et مَسْئُولِ plutôt que مَسْؤُولِ /masʔu:l/ ‘responsable’, mais aussi رُؤُوسِ plutôt que رُؤُوسِ ‘têtes’ avec hamza seule sur la ligne quand le graphe précédent ne se lie pas au suivant (Amīn et Tarzi, 1984, p. 311)⁸. Par contre, dans une apparente contradiction, elle préconise يَدُّوْنَ (avec une suite de deux و (ibid., p. 309) plutôt que le plus moderne يَدَّوْنَ (Badawi, Carter et Gully, 2016, p. 13) qui considère que l’ajout du suffixe du pluriel وَن ne transforme pas cette hamza finale en hamza médiane. Par ailleurs, concernant ce traitement de la hamza de troisième radicale, on la trouve seule sur la ligne dans le Coran وَيَدْرءُونَ بِالْحَسَنَةِ السَّيِّئَةِ ‘et repoussent le mal par le bien’ (13:22). Une suite de trois و est même possible avec le verbe يَسُوؤُونَ /jasu:ʔu:na/ ‘ils agissent mal’ (Ammar et Dichy, 2008, p. 188) (Reig, 1984, p. 162) mais ailleurs elle semble être évitée pour proposer les alternatives يَسُوؤُونَ يَسُوؤُونَ (Mace, 1999, p. 185).

La hamza s’avère donc un signe, dont la valeur phonologique est univoque, le “coup de glotte”, mais dont la réalisation écrite est plurielle, voire instable.

2.3 tā’ marbūṭa

La tā’ marbūṭa تاء مَرْبُوطَة, ‘tā’ liée’, est un graphe, toujours placé à la finale du mot, qui prend la forme du graphe هـ, surmonté de deux points. Il peut être lié à la lettre

⁸L’usage semble hésiter entre ces différentes orthographes comme une requête de ces mots entre guillemets sur le moteur de recherche Google l’indique avec (en milliers de résultats) un score de 23.800 pour سُئُونِ contre 24.400 pour سُؤُونِ, de 42.500 pour مَسْئُولِ contre 52.000 pour مَسْؤُولِ et de 355 pour رُؤُوسِ contre 16.300 pour رُؤُوسِ

précédente *فُرْشَة* /furʃa/ ‘brosse’ ou non, si la lettre précédente ne se lie pas : *فُرْشَاء* /‘furʃa:/ ‘brosse’. Il est obligatoirement précédé d’une consonne portant la fatha /a/, ou d’une ‘alif de voyelle longue /a:/. Il se prononce /t/ et porte une voyelle de déclinaison s’il est suivi d’un autre mot : *فُرْشَاءُ أُسْنَانٍ* /‘furʃa:tu ʔasna:n/ ‘brosse à dents’. Sinon, en fin de phrase, ou à la pause, il ne se prononce pas et le mot se termine par le son /a/ ou /a:/. Il sert généralement de marque du féminin (tā’ al-ta’nīl) des noms et des adjectifs : *زَوْج* /zawʒ/ ‘époux’ > *زَوْجَة* /zawʒa/ ‘épouse’, ou de discrétisation d’un nom collectif : *سَمَك* /samak/ ‘du poisson’ > *سَمَكَة* /samaka/ ‘un poisson’.

Ce signe a donc deux réalisations phonologiques, selon le contexte, d’où le non recours à la tā’ maftūḥa *تاء مفتوحة* (traditionnelle) *ت* /t/, sauf si un pronom clitique lui est affixé : *زَوْجَتِي* /zawʒatī/ ‘mon épouse’. Pour ce qui est de l’écriture, dans le Coran, cette finale était normalement orthographiée comme un hā’ هـ, la prononciation étant /ah/ ‘(devenue ultérieurement seulement /a/) à la pause (Putten, 2017, note 8) et /at/ avec déclinaison en liaison. La tā’ maftūḥa *ت* /t/ est aussi attestée dans le Coran pour cette terminaison (ibid., note 7). La prononciation /ah/ reste matérialisée dans la translittération traditionnelle qui utilise souvent un ⟨h⟩ (lequel ne se prononce plus, mais indique la présence d’une tā’ marbūṭa en arabe, comme dans cette translittération de *خليفة* ”khalifah” = Calife).

2.4 wāw et yā’

Chacun de ces deux graphes a deux valeurs phonologiques principales en arabe : voyelle longue ou semi-voyelle.

2.4.1 wāw

Le graphe wāw و peut transcrire la voyelle longue /u:/ : *نور* /nu:r/ ‘lumière’. Théoriquement, cela implique qu’il porte un sukūn et que la consonne précédente porte la voyelle courte ḍamma /u/ dont il est la prolongation *نُور*. En pratique, ces diacritiques apparaissent rarement.

Il peut transcrire la semi-voyelle (spirante labio-vélaire voisée) /w/ notamment pour la diphtongue /aw/ *نوم* /nawm/ ‘sommeil’.

Il peut aussi être support de hamza, selon les règles vues 40 *بؤس* /buʔs/ ‘malheur’.

On trouve un wāw qui ne se prononce pas dans l’orthographe du prénom *عمرو* /ʕamr/ ‘Amr’ pour le distinguer de *عمر* /ʕumar/ ‘Umar’.

2.4.2 yā’

Le graphe yā’ ي peut transcrire la voyelle longue /i:/ : *ريف* /ri:f/ ‘campagne’. Théoriquement, cela implique qu’il porte un sukūn et que la consonne précédente porte la voyelle courte kasra /i/ dont il est la prolongation *رِيف*. En pratique, ces diacritiques apparaissent rarement.

Il peut transcrire la semi-voyelle (spirante palatale voisée) /j/ notamment pour la diph-tongue /aj/ **يَيْت** /bajt/ ‘maison’. Cette semi- voyelle s’emploie aussi, sous une forme redou- blée, pour former le suffixe clitique de formation d’un adjectif nisba, c’est-à-dire dérivé d’un nom : **مَصْر** /miṣr/ ‘Égypte’ > **مَصْرِيّ** /miṣrijj/ ‘Égyptien’. C’est une suffixation très pro- ductive.

2.5 lām

Le graphe lām ل a des particularités d’écriture, notamment dans les suites lām ’alif et ’alif lām, cette dernière correspondant à l’article défini, lām al-ta’rīf.

2.5.1 lām ’alif

Lorsque la lettre lām ل est suivie d’une ’alif ا, les deux lettres sont systématiquement ligaturées dans une boucle : *ل > لا : لا /la:/ ‘non’, probablement pour rendre plus lisible cette suite de deux graphes formés chacun d’un trait verticale (la différence étant que lām se lie à la lettre suivante). On remarque que même des systèmes calligraphique visant la correspondance «une à une» des types d’imprimerie et des lettres, incluent cette ligature dans leur nombre limité de fontes, comme celui de l’Académie de la langue arabe du Caire de 1959 ou celui d’Aḥmad Laḥdar-Ġazal, lequel a dû rajouter une première fois la ligature lām ’alif à son «pur» système originel, et plus tard encore ses différentes déclinaisons لا, لا et لا (Haralambous, 1998). Ceci souligne l’exception scripturale que représente cette ligature.

2.5.2 lām al-ta’rīf

En arabe, l’article défini (لام التعريف - lām al-ta’rīf), est formé par la cliticisation en début de mot d’une ’alif support de la hamzat al-waṣl suivie du graphe lām ل **يَيْت** /bajt/ ‘(une) maison’ > **الْبَيْت** /al-bajt/ ‘la maison’. Le lām est lié dans l’écriture au mot qui le suit.

Comme indiqué page 39, la ’alif support de la hamzat al-waṣl ne se prononce pas, sauf en début de phrase. Après un autre mot, la liaison se fait par la voyelle finale de ce mot, ou par une voyelle épenthétique s’il se termine par une consonne. On peut ainsi dire que cette ’alif est la matérialisation à l’écrit de cette voyelle épenthétique potentielle.

Pour la prononciation du lām de l’article, il y a deux possibilités :

- le graphe lām se prononce /l/ quand le mot qui suit commence par un des phonèmes suivants : les bilabiales /m/ م et /b/ ب, la labiodentale /f/ ف, les glottales /ʔ/ أ et /h/ ه, les pharyngales /ħ/ ح et /ʕ/ ع, l’uvulaire /q/ ق, les vélaires ou post-vélaires /w/ و, /k/ ك, /x/ خ et /ɣ/ غ, les post-alvéolaires /dʒ/ ج ou /ʒ/ ج, la palatale /j/ ي (soit les consonnes des points d’ar- ticulation antérieurs ou postérieurs, les plus éloignés de /l/). Par exemple **القَمَر** /al-qamar/ ‘la lune’, d’où le nom de «lettres lunaires» données à ces graphes dans la grammaire tra- ditionnelle arabe. Dans les textes vocalisés, on écrit le diacritique sukūn au dessus du lām.

- le graphe lām ne se prononce pas et la première consonne du mot qui suit est redoublée (dans un phénomène d'assimilation), quand le mot qui suit commence par un des phonèmes suivants : les dentales /θ/ ث et /ð/ ذ, les alvéolaires /n/ ن, /r/ ر, /l/ ل (par nature), /t/ ت, /d/ د, /s/ س, /z/ ز plus les emphatiques de ces dernières /tˤ/ ط, /dˤ/ ض, /sˤ/ ص, /ðˤ/ ظ, la palato-alvéolaire /ʃ/ ش (soit les consonnes coronales, celles des points d'articulation médians, les plus proches de /l/). Par exemple الشَّمْسُ /aʃ-fams/ 'le soleil', d'où le nom de « lettres solaires » données à ces graphes dans la grammaire traditionnelle arabe. Dans les textes vocalisés, on écrit le lām sans diacritique et on place une šadda ʾ au dessus de la première consonne du mot.

La suite 'alif lām a donc de multiples réalisations phonologiques. Plutôt que de refléter systématiquement l'exacte prononciation des mots déterminés par l'article, on préfère maintenir l'unité de représentation de cet outil grammatical pour des raisons morphologiques.

Cet article est régulièrement précédé par un clitique affixé, comme par exemple la préposition كَ /ka/ 'comme' : كَالشَّمْسِ /kaʃ-fams/ 'comme le soleil'. Mais avec la préposition clitique لِ /li/ 'pour', la suite potentielle lām 'alif lām : *لِلشَّمْسِ est évitée en supprimant la 'alif support de la hamzat al-waṣl لِلشَّمْسِ /liʃ-fams/ 'pour le soleil'. Si, en plus, le mot commence par le graphe lām, la suite potentielle de trois lām *لِللَّيْلَةِ est évitée en n'écrivant que deux lām لِلَّيْلَةِ /li-l-lajla/ 'pour la nuit', créant ainsi une ambiguïté si la šadda disparaît (ce qui est souvent le cas) avec le sens لِلَّيْلَةِ /li-lajla/ 'pour une nuit'. L'écriture est donc ici simplifiée, au prix de la perte d'unité de représentation de l'article et d'une certaine ambiguïté sémantique.

2.6 la gémiation par šadda

En arabe, quand une consonne est immédiatement suivie de la même consonne, sans voyelle intermédiaire, une seule lettre est écrite, et le signe diacritique šadda ʾ peut être écrit sur cette consonne. Cela correspond au doublement de la consonne, ou gémiation, appelé tašdīd (renforcement) par les grammairiens arabes. C'est le phénomène que l'on observe dans la différenciation des verbes de première forme, comme عَلَّمَ /ʕalima/ 'il sut', de ceux de forme ii عَلَّمَ /ʕallama/ 'il enseigna' et de forme v تَعَلَّمَ /taʕallama/ 'il apprit'. La distinction entre les deux premiers n'apparaissant pas à l'écrit si la šadda n'est pas indiquée, elle se fera par induction à partir du contexte.

Cette gémiation s'applique aussi aux graphes représentant les suffixes de conjugaison, pour les verbes qui se terminent par une consonne similaire. Ainsi, on trouve سَكَّتُ /sakattu/ 'je me tus', à comparer avec سَكَّنْتُ /sakantu/ 'je me suis calmé', ou inversement سَكَّنا /sakanna/ 'nous nous sommes calmés', à comparer avec سَكَّنَّا /sakatna/ 'nous nous tûmes'. Dans ce cas, la gémiation abolit la distinction entre la racine et le suffixe, ce qui permet au scripteur d'économiser l'écriture d'une lettre, mais requiert du lecteur un effort supplémentaire de reconnaissance de la double valeur de la consonne finale, surtout si la šadda n'est pas écrite.

La gémation est également utilisée pour des assimilations telles que *مِنْ /min ma:/ 'de qui', *عَنْ /ʕan man/ 'de qui', *إِنْ /ʔin la:/ 'sauf', ...

Inversement, malgré l'assimilation du phonème /l/ 'de l'article défini (cf. supra 2.5.2, page 43) avec les phonèmes transcrits par les lettres « solaires », le graphe lām est conservé à l'écrit : الشَّمْسُ /aʃ-fams/ 'le soleil'. L'orthographe arabe, en gardant le graphe lām, qui n'est pas prononcé, privilégie donc ici la morphologie, au détriment de la représentation phonologique stricte.

Wright parle de « tašdīd euphonique » pour les deux cas d'assimilation ci-dessus (par opposition au « tašdīd nécessaire » des formes verbales citées), et complète avec d'autres exemples dont أَرَدْتُ /ʔaradta/ 'je voulais' > أَرَدْتُ /ʔaratta/ 'tu voulais'. Il précise cependant :

Many grammarians, however, reject this kind of assimilation altogether, and rightly, because the absorption of a strong radical consonant, such as د [...], by a weaker servile letter, like ت, is an unnatural mutilation of an essential part of the word. (Wright, 1967, p. 16)

De nombreux grammairiens rejettent cependant totalement ce type d'assimilation, et à juste titre, car l'absorption d'une consonne radicale forte, comme د [...], par une lettre servile plus faible, comme ت, est une mutilation contre nature d'une partie essentielle du mot.

On constate ainsi un autre exemple de tension entre une écriture basée sur la prononciation perçue des mots et une écriture respectueuse de la morphologie de ceux-ci.

2.7 Autres assimilations

2.7.1 Les verbes de forme viii

La forme viii des verbes se caractérise par l'infexion de /t/ après la première radicale des verbes trilitères. Ainsi, à partir de la première forme غَسَلَ /ḡasala/ 'il lava' on obtient par cette règle la forme viii اِغْتَسَلَ /iḡtasala/ 'il se lava'. Pour les verbes dont la première radicale est /t/, la graphie fusionnera les deux ت correspondants sous une éventuelle šadda : تَبَعَ /tabiʕa/ 'il suivit' > اِتَّبَعَ /ittabaʕa/ 'il poursuivit'. Le contact étroit entre certaines premières radicales et cet infixe consonne occlusive alvéolaire sourde /t/ entraîne des phénomènes d'assimilation que l'arabe note graphiquement.

Si cette première radicale est sa jumelle voisée /d/, l'infixe s'assimile en /d/ : دَعَمَ /daʕa-ma/ 'il soutint' > اَدَّعَمَ /ʔiddaʕama/ 'il s'appuya [sur]'. De même avec la fricative dentale voisée /ð/ : ذَخَرَ /ðaxara/ 'il épargna' > اِدَّخَرَ /ʔiddaxara/ 'il stocka'. La fricative alvéolaire voisée /z/ subsiste mais sonorise l'infixe en /d/ : زَهَرَ /zahara/ 'il brilla' > اِزْدَهَرَ /ʔizdahara/ 'il prospéra'. Avec la fricative dentale sourde /θ/, l'infixe s'assimile en /θ/ : ثَأَّرَ /θaʔara/ 'il vengea' > اِثَّأَّرَ /ʔiθθaʔara/ 'il se vengea'.

Si cette première radicale est une des consonnes emphatiques, l'infixe /t/ devient /tʕ/ : اِطَّلَعَ /tʕalaʕa/ 'il monta' > اِطَّطَّلَعَ /ʔitʕtʕalaʕa/ 'il s'informa', صَدَّمَ /sʕadama/ 'il heurta' > اِصَّطَدَّمَ /ʔisʕtʕadama/ 'il heurta'.

/(?̣)isʔtʕadama/ ‘il entra en collision’, ضَرَبَ /dʕaraba/ ‘il frappa’ > اِضْطَرَبَ /(?̣)idʕtʕaraba/ ‘il fut troublé’.

Ces assimilations induites par la proximité de l’infixe des verbes de forme viii avec certains types de premières radicales sont donc effectivement retranscrites dans la graphie de l’arabe.

2.7.2 assimilations non transcrites

Les locuteurs de l’arabe sont conscients d’autres phénomènes d’assimilation, lesquels ne sont pourtant pas transcrits dans la graphie. Ainsi la rencontre entre la consonne occlusive bilabiale voisée /b/ et la nasale alvéolaire voisée /n/ a tendance à transformer celle-ci en bilabiale /m/ comme dans les exemples suivants (Al-Sāwī, 2017, p. 24) :

- أَنبِيَاءَ ?/ʔanbiʔa:ʔ/ ‘prophètes’ > * أَمْبِيَاءَ ?/ʔambiʔa:ʔ/
- أَبْنَاءَ ?/ʔanba:ʔ/ ‘bulletins’ > * أَمْبَاءَ ?/ʔamba:ʔ/
- إِنْهَارَ ?/ʔinbiha:r/ ‘éblouissement’ > * إِمْبَاهَارَ ?/ʔimbiha:r/

Ainsi, l’orthographe arabe accepte de transcrire certains phénomènes d’assimilation, rendant ainsi plus difficile l’opération de dérivation que le lecteur doit effectuer à partir de la racine. Dans d’autres cas, l’aspect morphologique est respecté, même si le même phénomène phonétique justifierait une modification de la graphie.

2.8 Conclusion

L’écriture arabe est un moyen de transcription de la fuṣḥā qui a prouvé son efficacité tout au long de l’histoire, notamment par son grand respect de la règle d’univocité de la plupart de ses graphes. Cette histoire a cependant laissé des traces, certains graphes ayant changé de valeur, d’autres s’étant ajoutés, ils sont parfois devenus équivoques ou alors certains phonèmes sont représentés par différents graphes, selon les contextes. Le choix est généralement fait de privilégier une représentation fidèle de ce qui est considéré comme la prononciation pertinente, mais une orthographe morphologique subsiste parfois.

De plus, l’absence d’indications des voyelles courtes et des autres phénomènes marqués par des diacritiques, s’il facilite la tâche du scripteur, oblige cependant le lecteur à faire émerger le sens en l’amenant à compléter les informations manquantes, notamment en identifiant ce qui ressort de la racine et ce qui est affixe. L’emploi des graphes arabes pour l’arabe égyptien soulèvera les mêmes problématiques, mais s’y ajoutera une plus grande disparité entre la prononciation de la fuṣḥā, influencée par son écriture, et celle de la ‘ammiyya, dont les règles sont parfois très différentes.

Chapitre 3

Naissance et enjeux de Wikipédia maŕri

3.1 Présentation

Wikipédia est une encyclopédie en ligne fondée par Jimmy Wales et Larry Sanger en janvier 2001 (Crovitz et Smoot, 2009, p. 91). Le terme Wikipédia est formé sur le modèle de la centaurisation, c'est-à-dire «l'écrasement de deux mots» (Miannay, 1972, p. 15) en l'occurrence le mot anglais *encyclopedia*, aphérésé et le mot hawaïen *wiki* signifiant «rapide», qui prend ici le sens de «rapidement et facilement modifiable par l'utilisateur» (Hasan, 2012, p. 11). Il s'agit donc d'une encyclopédie collaborative et d'une œuvre libre, c'est-à-dire que chacun est libre de l'amender et de la rediffuser (Crovitz et Smoot, 2009, p. 92).

Lancée en janvier 2001 en tant qu'encyclopédie de langue anglaise (code en), des éditions dans d'autres langues ont rapidement été créées : en allemand (de), en catalan (ca) et en français (fr) en mars suivant. En avril 2025, des articles Wikipédia avaient été créés dans 355 éditions, dont 341 actives et 14 fermées (Wikimedia, 2025b). Pour être éligible, une langue doit remplir certaines conditions définies dans la page en anglais faisant foi et intitulée “Language proposal policy” traduites de la manière suivante dans son équivalent en langue française intitulée “Règlement des propositions de langues” :

Conditions d'éligibilité

1. La demande concerne l'ouverture de l'édition d'une nouvelle langue d'un projet Wikimedia existant qui n'existe pas déjà [...].
2. La langue doit généralement posséder un code ISO-639 1–3 valide [...].
3. La langue doit être suffisamment unique pour ne pas pouvoir coexister sur un wiki plus général. Dans la plupart des cas, cela exclut les dialectes régionaux et les différentes formes écrites de la même langue. [...]
4. La langue de la proposition a un nombre suffisant de locuteurs courants pour former une communauté de contributeurs viable et avoir un public pour le contenu. [...] (Wikimedia, 2025d) ¹.

¹Requisites for eligibility 1. The proposal is to open a new language edition of an existing Wikimedia project

L'édition de Wikipédia en langue arabe (code ar) a été lancée en juillet 2003. Ont aussi leur propre édition, le maltais (mt) depuis septembre 2004 (langue dérivée de l'arabe), l'arabe égyptien (arz), depuis novembre 2008 et l'arabe marocain (ary), depuis juillet 2020. Ces deux dernières éditions ont réussi à remplir les conditions d'éligibilité, y compris la troisième qui pourtant exclut les "dialectes". Cela révèle l'arbitraire de ce terme illustré par l'aphorisme popularisé par le linguiste yiddish Max Weinreich (1894-1969) : "Une langue est un dialecte avec une armée et une marine".

Wikipédia en arabe égyptien (Wikipédia maṣri² selon l'appellation propre de l'édition) a démarré en tant que projet en développement le 2 avril 2008 dans l'incubateur Wikimedia. Au terme de la discussion, la demande pour l'ajout de l'arabe égyptien en tant que nouvelle langue sur Wikipédia a été acceptée en juillet et l'annonce a en été faite le 17, premier jour de Wikimania³ 2008, à Alexandrie. Le 24 novembre 2008, Wikipédia en arabe égyptien a été officiellement lancé et les articles de l'incubateur ont été transférés vers le nouveau domaine (Wikipédia, 2025d).

Depuis lors, le nombre d'article n'a cessé d'augmenter. Panović (2010, p. 94) note qu'ils étaient au nombre de 6.000 au bout de deux ans d'existence.

Année	2009	2010	2011	2012	2013	2014	2015	2016
Articles	3.952	2.708	1.344	1.094	1.698	1.893	1.028	1.646
Année	2017	2018	2019	2020	2021	2022	2023	2024
Articles	1.515	2.303	19.770	1.143.950	328.316	103.968	5.244	5.110

Tab. 3.1: Nombre annuel d'articles créés sur Wikipédia maṣri

On observe dans le tableau 3.1 un enthousiasme du début avec près de 4.000 nouveaux articles en 2009 et plus de 2.700 en 2010, mais ensuite la moyenne annuelle de création de pages de contenu s'établit entre 1.000 et 2.000 avec une légère accélération en 2018 avec 2.300 articles. Une explosion du nombre de créations de pages de contenu a lieu à partir de 2019 pour atteindre plus d'un million en 2020 (avec une pointe à plus de 250.000 en juin). Le rythme redescend ensuite et se situe à une moyenne d'environ 5.000 nouveaux articles par an en 2023 et 2024.

L'explication de ce pic de créativité se trouve dans le livre accessible par un lien situé sur la page de l'utilisateur HitomiAkane. Ce livre publié en 2021 s'intitule *How I wrote a million Wikipedia articles* 'Comment j'ai écrit un million d'articles sur Wikipédia'. C'est le récit autobiographique de Maher Asaad Baker, un Syrien réfugié en Égypte pour fuir la guerre civile dans son pays. Il critique le sectarisme qu'il a ressenti en participant à Wikipédia arabe, notamment de la part des utilisateurs qui affichaient leur hostilité envers Wikipédia maṣri, ce qui l'a poussé à participer à ce dernier projet à partir de l'été 2019. Il

that does not already exist [...]. 2. The language must ordinarily have a valid ISO 639 1–3 code [...]. 3. The language must be sufficiently unique that it could not coexist on a more general wiki. In most cases, this excludes regional dialects and different written forms of the same language. [...] 4. The language of the proposal has a sufficient number of fluent users to form a viable contributor community and an audience for the content. [...]' (Wikimedia, 2025a)

²En arabe égyptien, l'adjectif maṣri مصري signifie "égyptien" et est dérivé de maṣr مصر qui signifie "Égypte", la voyelle /a/ remplaçant le /i/ de l'arabe littéral.

³Wikimania est une conférence internationale rassemblant chaque année depuis 2005 des personnes intéressées par les projets hébergés par la fondation Wikimedia et le mouvement Wikimedia en général.

explique comment il a écrit un programme pour créer des articles à la chaîne à partir des données disponibles sur Wikidata, une "base de connaissances libre et ouverte" qui "est le dépôt central de données structurées de Wikimedia pour ses projets frères, dont Wikipédia [...]" (Wikidata, 2023). Il combine ainsi variables et valeurs pour écrire de courts articles sur le même type de sujets, des joueurs de football par exemple, avec une ou deux phrases types dont seules les informations concernant chaque joueur changent (nom, nationalité, club, ...). Il décrit en détail la méthode qu'il a suivie et qui lui a permis de faire bondir le nombre d'articles de Wikipédia en arabe égyptien (Baker, 2021).

Au 14 avril 2025, ces articles sont au nombre de 1.626.778, ce qui place Wikipédia en arabe égyptien en onzième position selon ce critère parmi les éditions Wikipédias des différentes langues, la première place étant occupée par l'anglais avec 6.981.176 articles, le français étant en quatrième position avec 2.677.456 articles et l'arabe en dix-septième position avec 1.258.615 articles (Wikimedia, 2025b).

Cependant, Panović témoigne déjà en 2010 comment beaucoup des articles sont des «ébauches», selon la terminologie de Wikipédia, c'est-à-dire des articles très courts, en attente de développement. Il décrit une stratégie de la part des Wikipédiens mašri consistant à débiter de nouveaux articles juste pour en augmenter le nombre, quitte à les développer plus tard (Panović, 2010, p. 94-95). Pour nuancer l'évaluation largement quantitative créée par le classement selon le nombre d'articles, un autre classement a été créé, basé sur la profondeur d'édition des articles, «l'un des nombreux indicateurs bruts de la qualité collaborative de l'encyclopédie, reflétant la fréquence à laquelle les articles sont mis à jour». Elle est calculée selon la formule suivante⁴ :

$$Profondeur = \left(\frac{\text{Éditions}}{\text{Total}} \right) \cdot \left(\frac{\text{Non Articles}}{\text{Articles}} \right)^2$$

Selon ce classement appliqué aux éditions avec plus de 10.000 articles, au 14 avril 2025, l'édition anglaise occupe la première place, avec un indice de profondeur de 1307,18, l'édition arabe est sixième (283,29), la française septième (269,82) et l'édition arabe égyptien est au 159^e rang (sur 162) avec un indice de 0,55 (Wikimedia, 2025e). L'observation faite par Panović en 2010 semble donc toujours valable, les rédacteurs de Wikipédia en arabe égyptien privilégiant le nombre d'articles sur leur profondeur, ce qui leur a permis de figurer parmi le premier cercle des 15 éditions ayant plus de 100.000 articles sur la page principale de l'édition anglaise de Wikipedia jusqu'en 2024 (Wikipédia, 2024c).

La comparaison avec des éditions comptant un nombre similaire de locuteurs au niveau de vie comparable permet de nuancer l'image comme il apparaît dans le tableau 3.2⁵.

Wikipédia mašri se situe en dessous de la moyenne sans que le nombre de ses utilisateurs, utilisateurs actifs et administrateurs soit insignifiant (Wikimedia, 2025c). Il faut par ailleurs tenir compte de la concurrence que Wikipédia en langue arabe constitue spécifiquement en Égypte pour pondérer ces indicateurs.

⁴«Non-Articles réfère aux pages de discussions, aux redirections et aux pages utilisateurs, et Total réfère à Non-Articles + Articles.» (Wikimedia, 2025e)

⁵Le nombre de locuteurs inclut ceux de langue première et seconde. Source principale : <https://www.ethnologue.com/> et Wikipédia. U. = utilisateurs, L. = locuteurs, Admins = administrateurs.

Langue	Locuteurs	Articles	Profondeur	U. / Loc.	Utilisateurs	U. Actifs	Admins
arabe égyptien	53.990.000	1.626.778	0,55	0,0048	259.534	195	7
pachtou	60.000.000	20.504	31,3	0,0005	35.538	44	3
ourdou	63.948.800	220.257	246,6	0,0030	192.837	259	8
tamil	68.763.360	173.397	41,9	0,0035	242.692	297	32
turc	70.805.930	632.791	189,3	0,0233	1.655.364	2.661	23
marathi	71.780.660	99.775	39,2	0,0023	169.609	177	10
persan	72.279.286	1.035.958	152,8	0,0191	1.383.694	4.962	35
vietnamien	73.600.000	1.293.899	745,1	0,0133	985.756	1.28	15
télougou	74.049.000	111.637	70,6	0,0018	136.168	216	11
wu	80.000.000	45.320	0,9	0,0014	118.466	73	3
javanais	84.308.740	74.430	20,5	0,0007	66.038	96	5
haoussa	100.000.000	56.857	2,2	0,0003	34.631	137	9
moyenne	72.793.815	440.027	128,4	0,0062	440.027	770	13

Tab. 3.2: Comparaison de différentes éditions de Wikipédia

Si l'on regarde la répartition des pages vues par pays (demandes mensuelles ou moyennes quotidiennes, pour septembre 2018, les derniers chiffres trouvés), on observe que 59,3% de celles vues en Égypte relèvent de Wikipédia en arabe, 33,3% en anglais, 2,9% en russe et 2,1% en arabe égyptien (Wikimedia, 2018b). Parallèlement, la répartition des pages vues par édition révèle que pour Wikipédia maṣri, 34,9% l'ont été à partir de l'Égypte, 12,2% à partir des États-Unis, 10,9% à partir de l'Arabie Saoudite, le reste l'étant essentiellement à partir des autres pays arabes (Wikimedia, 2024). L'article en anglais sur cette édition de l'encyclopédie indique "En décembre 2022, c'était la troisième édition Wikipédia la plus visitée en Égypte avec 2 millions de pages vues. Elle se classait derrière la Wikipédia arabe (43 millions) et la Wikipédia anglaise (18 millions)."⁶ (Wikipédia, 2025d). Mais ces chiffres n'ont pas pu être confirmés par la consultation du site Wikimedia (2025f) indiqué en référence.

Le décompte des appareils uniques qui visitent le projet par mois en moyenne montre une progression régulière chaque année d'à peine 700.000 en 2016 à près de 3 millions en 2024, avec un pic à plus de 6.336.000 en janvier 2025 (pour Wikipédia en arabe, l'évolution va d'environ 20 millions à environ 48 millions sur la même période). L'article le plus vu est *Lista tartīb il-mumassilāt il-būrṇūgrāfīyāt ḥasab is-sinn min il-'akbar lil-'aṣḡar* 'Liste des actrices pornographiques par âge, de la plus âgée à la plus jeune', en 2016 (Nabulssi-Maselbas, 2023, p. 40) et en 2024, avec près de 40.000 vues mensuelles en moyenne (Wikimedia, 2025f). Cet intérêt pour la pornographie reflète la réalité d'internet, le site *pornhub.com* étant le huitième site le plus visité en février 2025, derrière, *google.com*, *youtube.com* et *facebook.com*, *instagram.com*, *wikipedia.org*, *reddit.com* et *bing.com* et devant *chatgpt.com* et *x.com* (Semrush, 2025). En 2024, outre la page principale, l'article consacré à la lettre *س* est bien classé, ainsi que *عيد الام* 'la fête des mères', ce qui reflète la diversité des intérêts des utilisateurs. Le nombre de rédacteurs actifs (avec au moins 5 modifications par mois) a évolué aussi, passant d'une fourchette de 10 à 30

⁶"In December 2022, it was the third most visited language Wikipedia in Egypt with 2 million page views. It ranked below the Arabic Wikipedia (43 million) and the English Wikipedia (18 million)."

jusqu'en 2020, à une fourchette de 20 à 46 depuis (Wikimedia, 2025f).

Wikipédia en arabe égyptien reste donc une aventure qui rassemble un ensemble de passionnés, dont l'impact sur la société égyptienne reste pour l'instant limité, quoique non négligeable. Il s'agit de fait d'une sorte de laboratoire où des locuteurs de l'arabe égyptien, quelles que soient leurs motivations, s'appliquent à écrire cette langue sans norme orthographique, dans des domaines inédits, tels que peuvent être ceux d'une encyclopédie. Pour cela, l'étude des choix linguistiques et graphématiques qui y sont faits révèle les lignes de force issues de la confrontation entre fuṣḥā et 'āmmiyya. Le développement du nombre de ses utilisateurs et rédacteurs laisse penser que son impact va augmenter à l'avenir, et donc que les options privilégiées en son sein sont appelées à voir leur influence croître.

3.2 Débat autour du projet

Le projet a été véritablement lancé avec la création sur Wikipédia le 30 mars 2008 d'une page de demande pour l'ajout de l'arabe égyptien en tant que nouvelle langue. L'initiateur est un Wikipédien portant le nom d'utilisateur Ghaly, lequel est l'un des principaux champions de la cause. Conformément à la procédure Wikipédia, une discussion y a eu lieu, essentiellement d'avril à juillet, où des contributeurs ont écrit, en anglais, les arguments pour et contre (Wikimedia, 2018a). Ils sont traduits, résumés et réorganisés dans les tableaux 3.3, 3.4 et 3.5.

Concurrence avec Wikipédia arabe	
Arguments contre	Arguments pour
Cela nuira certainement à Wikipédia arabe, car cela ne fera que puiser dans ses ressources.	Wikipédia arabe n'est pas censé avoir le monopole de l'information et de son accès.
Dans l'article sur Venise, il n'y a pas de différences majeures entre l'article arabe classique et l'article postulé de Masri en termes de mots. En revanche, le même sujet en langue persane, qui utilise l'alphabet arabe dans l'écriture, ne pourrait être lu par quiconque originaire des pays arabes sans connaissance préalable de cette langue.	Le nom de l'article en arabe égyptien est différent du nom en arabe car le nom moderne de la ville est utilisé en arabe égyptien. Merci aussi de reconnaître qu'il y a 11 mots différents (sur 120).
Ce type de proposition est motivé par des raisons nationalistes et politiques.	Wikipédia devrait-il refuser ce projet au nom du nassérisme ou du panarabisme ?

Tab. 3.3: Arguments sur la concurrence avec Wikipédia arabe

Statut de l'arabe égyptien	
Arguments contre	Arguments pour
L'arabe égyptien est un dialecte ou un groupe dialectal, il n'est pas suffisamment différent pour nécessiter un Wikipédia.	L'arabe égyptien une langue parlée par plus de 70 millions de personnes ⁷ et l'ISO 639-3 lui attribue le code arz.
Le gouvernement égyptien affirme que la langue officielle de l'État est l'arabe.	Les porte-paroles du gouvernement égyptien utilisent l'arabe égyptien pour parler au public.
L'Égypte compte de nombreux dialectes divers dans tout le pays. Comment ceux hors du Caire vont-ils comprendre les Cairotes qui écrivent le qāf avec une hamza ? Comment les Cairotes vont-ils comprendre ceux du Ṣa'īd s'ils transcrivent [g] par le qāf ? Il n'existe pas de règles de grammaire claires pour l'arabe égyptien, ni pour l'écriture des mots.	Écris comme tu le souhaites. La manière de s'exprimer du Caire est la plus connue pour écrire l'arabe égyptien. Écrivez en Baḥrawī, en Ṣa'īdi ⁸ , ... L'important est que tu écrives d'une manière que les Égyptiens puissent comprendre.
L'écriture de l'arabe égyptien est une écriture arabe mais pleine d'erreurs.	Ce que tu considères comme une erreur est une différence.
Voulons-nous avoir 22 autres wikipédias en dialecte arabe ? Plusieurs wikipédias pour la même langue mais différents systèmes d'écriture sont une perte d'énergie.	L'arabe égyptien possède son propre vocabulaire et devrait être considéré comme une langue.
Accepteriez-vous d'avoir un Wikipédia en anglais américain, un Wikipédia en anglais australien, un Wikipédia en anglais britannique, un Wikipédia en anglais néo-zélandais, un Wikipédia en anglais hindi ?”	La langue utilisée par le wikipédia anglais peut facilement être comprise par tout lecteur dont la langue maternelle est l'anglais, alors que ce n'est pas le cas dans le wikipédia arabe.

Tab. 3.4: Arguments sur le statut de l'arabe égyptien

⁷La Banque mondiale évalue à environ 84 millions la population égyptienne en 2008 (contre 111 millions en 2022) (WorldBank, 2024), l'arabe égyptien étant attribuable à un nombre variable d'Égyptiens, selon qu'on limite sa définition au parler du Caire, ou qu'on englobe d'autres régiolectes plus ou moins proches.

⁸En Égypte, le terme « baḥrawī » réfère à ce qui est relatif à la Basse-Égypte, c'est-à-dire le delta, face à la mer (baḥr) et le terme « ṣa'īdi » réfère à ce qui est relatif à la Haute-Égypte, soit le sud du pays considéré comme l'amont, les hautes terres (ṣa'īd).

Richesse de l'arabe égyptien	
Arguments contre	Arguments pour
Vous n'avez pas encore prouvé que ce que vous appelez «l'arabe égyptien» est une langue utilisée pour la littérature écrite, notamment scientifique et encyclopédique, une poignée de romans et de poèmes ne comptent pas ici. Il n'y a aucune ressource d'information respectée ou bien connue (comme un livre ou un article dans un site d'actualités) écrite en arabe égyptien.	Lorsqu'un médecin, un chirurgien, un dentiste ou un pharmacien s'adresse à un patient en Égypte, il utilise des termes développés par les enseignants de ces professionnels au fil des années de communication avec les patients, ce n'est pas du jargon mais de simples expressions en arabe égyptien.
Depuis quand les Égyptiens ont-ils leur propre jargon médical qui diffère de ce qu'il appelle «l'arabe classique» ?	C'est le langage qu'un médecin utilisera pour décrire une maladie à un patient, qu'un enseignant utilisera pour expliquer une équation, qu'un manager utilisera pour guider ses collègues dans la vie de tous les jours ;
Exprimer des idées complexes et abstraites et des principes scientifiques précis nécessite une grande discipline dans la forme sous laquelle ces idées sont exprimées. Continuez à parler à une personne en langage bébé et son cerveau restera celui d'un bébé, la superficialité est un danger.	Des informations de référence rédigées dans un langage plus familier satisferaient bien mieux au mandat de Wikimedia de rendre l'information accessible au plus grand nombre possible de personnes.
Les personnes analphabètes ne pourront pas lire ce qui est écrit dans ce Wikipédia en arabe égyptien, les gens instruits liront le texte de Wikipédia en arabe égyptien, mais ils liront également le texte de Wikipédia en arabe qui existe actuellement.	Qu'en est-il des Égyptiens qui ne parlent pas bien l'arabe - et ils constituent la majorité -, ou simplement des Égyptiens qui revendiquent le droit et la liberté d'utiliser leur propre langue ?
Les étudiants égyptiens à l'école et à l'université apprennent en arabe de Wikipédia arabe, la langue dans laquelle sont écrits les livres qu'ils lisent.	Mais les enseignants utilisent l'arabe égyptien lorsqu'ils expliquent les cours.

Tab. 3.5: Arguments sur la richesse de l'arabe égyptien

Apparaissent dans cette discussion beaucoup des arguments utilisés traditionnellement dans le débat sur la relation entre fuṣḥā et ʿammiyya. Les opposants à la reconnaissance de cette dernière s'interrogent notamment sur la réalité de l'existence même de l'arabe égyptien en tant que langue et mettent en avant l'indigence ressentie de son corpus écrit, l'absence de règles d'orthographe, son incapacité perçue à traiter de sujets complexes, la menace ressentie que son institutionnalisation fait peser sur l'identité arabe

telle que perçue. Les défenseurs du projet revendiquent la liberté de s'exprimer dans ce qu'ils considèrent comme leur langue nationale, la proximité avec les milieux populaires, la prise en compte pragmatique d'une situation diglossique vécue au quotidien.

Dès le 9 avril 2008, l'utilisateur Ghaly crée une page en arabe égyptien qui sera la politique officielle du projet pour la « méthode d'écriture » *طريقة الكتابة*.

Le recours à l'écriture pour la *‘āmmiyya* est justifié dès le début :

طريقة الكتابة في ويكيبيديا مصرى النسخة المصرية Wikipedia مكتوبه باللغة المصريه الحديثه الى يتكلمها المصريين و مكتوبه زى ما بيكتبوها في جواباتهم لبعضهم و في حياتهم اليوميه. المصريين بيكتبو المصري في الروايات, و المسرحيات, و القصائد الشعبيه و كان في المجلات الكوميدي و الاعلانات و شوية جرائين.

La méthode d'écriture dans Wikipédia en égyptien, la version égyptienne de Wikipédia, est écrite dans la langue égyptienne moderne parlée par les Égyptiens et est écrite comme ils l'écrivent dans leurs lettres entre eux et dans leur vie quotidienne. Les Égyptiens écrivent en égyptien dans des romans, des pièces de théâtre et des poèmes populaires, ainsi que dans des magazines de bandes dessinées, des publicités et quelques journaux. (Ghaly, le 28/07/2008)

Cependant, aucune référence précise n'est donnée, bien qu'elles soient relativement nombreuses, comme exposé dans la section 1.2.

La définition qui est faite de l'arabe égyptien dans la phrase d'introduction à cette page de recommandations est pratiquement tautologique puisqu'il s'agit de « La langue égyptienne moderne parlée par les Égyptiens ». De même, le premier conseil donné pour l'écriture est centré autour de cette égyptianité :

مكن يتكتب المصري بأى طريقة طالما أى مصرى ممكن يفهمها.

L'égyptien peut s'écrire de n'importe quelle manière tant que n'importe quel Égyptien peut le comprendre.

Si la question de la diversité des régiolectes en Égypte est abordée, ce n'est que pour encourager à un standard :

اكتب بالشكل الى يعجبك, بس القاهريه هيه الطريقه المشهوره لكتابة الكلام المصري. المهم يفهمها المصريين.

Écris comme tu le souhaites, mais le [parler] du Caire est la manière connue d'écrire l'égyptien, l'important est que les Égyptiens comprennent. (Ghaly, 9/4/2008)

La problématique de la hiérarchisation des différents régiolectes égyptiens, en termes de prestige et d'accès aux médias, est rapidement escamotée par la référence à un peuple égyptien qui est présenté succinctement comme étant uniforme (Wikipédia, 2025c).

L'appellation choisie pour la langue utilisée par le projet évolue en début de projet. Tout d'abord le terme utilisé par Ghaly sur cette page officielle est dès le 10/4 al-luğa al-*‘āmmiyya* al-*maṣriyya* *المصريه العاميه المصريه* ⁹ 'la langue familière égyptienne', avec un lien vers un article portant ce titre, puis, le 24/4, le titre de cet article devient *kalām maṣri* *كلام مصري* 'parler égyptien'. Le 22/01/2009, al-*‘āmmiyya* disparaît de l'appellation. Finalement, le 17/7/2010, Ghaly modifie l'appellation et choisit al-luğa al-*maṣriyya* al-*ḥadīsa*

⁹Dans ma transcription des titres ou citations de pages Wikipédia en arabe égyptien, j'appliquerai les principes de prononciation de l'arabe égyptien tels que décrits par Badawi et Hinds (1986).

اللغة المصرية الحديثة, c'est-à-dire 'la langue égyptienne moderne', raccourci dans le menu linguistique de Wikipédia en maṣri مصري, l'inscrivant ainsi dans un continuum idéalisé avec la langue égyptienne ancienne (hiéroglyphique) et effaçant son statut de langue du commun ('āmmiyya) voire son lien avec l'arabe. La différenciation d'avec ce dernier se révèle être un enjeu essentiel (Panović, 2010, p. 97-100), même si cela ne va pas jusqu'à abandonner l'abjad. Le lien hypertexte de l'appellation اللغة المصرية الحديثة mène vers un article Wikipédia maṣri qui porte ce même titre, et dont les équivalents sont en français «Arabe égyptien» et en arabe littéral لهجة مصرية lahǧa miṣriyya, c'est-à-dire «dialecte égyptien», ce qui illustre la diversité des points de vue concernant cet objet linguistique (Wikipédia, 2024a).

3.3 Élaboration de recommandations orthographiques

En mai 2008 une page Niqāš wīkībīdyā : tarīqat al-kitāba نقاش ويكيبيديا:طريقة الكتابة 'discussion Wikipédia : méthode d'écriture' est créée par Ghaly, la plupart des contributions étant en arabe égyptien, certaines en arabe littéral, notamment celles remettant en cause le bien-fondé du projet tout entier, et quelques-unes en anglais. Des problèmes d'écritures en arabe égyptien sont soulevés qui donnent lieu à des discussions, souvent succinctes.

3.3.1 Suggestions d'écritures en alphabet latin

Parmi les suggestions, il y a celles consistant à ne pas écrire avec les graphes arabes. Certains suggèrent d'utiliser الفرانكو al-frānkō (parfois الفرانكو ارب al-frānkō 'arab ou الانجلو al-'anglō 'arab), qui s'avère être une sorte d'arabizi avec utilisation des lettres latines pour les sons correspondants en arabe, et des chiffres, voire des digrammes, pour les autres. Certaines des contributions sont d'ailleurs écrites dans ce style. Ses adeptes mettent en avant la facilité d'écriture sur un clavier d'ordinateur ou de téléphone portable ne comportant pas les graphes arabes, tout en reconnaissant la subjectivité des règles d'orthographe appliquées, influencée notamment par la langue étrangère (anglaise ou française essentiellement) prise comme référence pour la transcription.

L'utilisateur MatthewS suggère une écriture «phonétique» avec des lettres latines éventuellement munies de diacritiques. Un tableau de correspondances est élaboré, les graphes étant classés dans l'ordre de l'alphabet latin. Les symboles choisis évolueront sensiblement au fil du temps (Wikipédia, 2021). Des exemples d'articles sont rédigés suivant différents systèmes de transcription (Wikipédia, 2025b). Le but est d'écrire l'arabe égyptien en représentant la prononciation de manière plus précise que l'abjad arabe. Cette méthode est même présentée comme une «libération de la langue égyptienne des lettres arabes incapables de représenter le discours égyptien»¹⁰ (archive 1). Les objections mises en avant sont la difficulté pour les Égyptiens de maîtriser ce système, auquel ils ne sont pas habitués, ainsi que la radicalité d'un tel choix d'écriture. De fait, les adversaires du projet Wikipédia maṣri brandissent régulièrement cette perspective d'écriture comme un argument massue visant à dénoncer une soumission à l'Occident et une menace contre la culture arabe et l'islam (Wikipédia, 2020a).

¹⁰ 'كتابة دقيقة و فونيتيك, تحرير للغة المصرية من الحروف العربية العاجزة عن تمثيل الكلام المصري'

Ces suggestions d'écritures en alphabet latin ne seront donc pas retenues. Il y a eu aussi des suggestions d'écriture avec l'alphabet grec, notamment parce que c'est de celui-ci qu'est issu l'alphabet copte, et même un tableau de translittération des lettres arabes avec l'alphabet cyrillique, mais le choix sera fait d'écrire l'arabe égyptien avec les graphes arabes.

3.3.2 Emploi des graphes arabes pour l'écriture de l'arabe égyptien

L'abjad arabe étant adapté à l'écriture de la fuṣḥā, les enjeux de son emploi pour l'écriture de la 'āmmiyya sont inévitablement évoqués par les porteurs du projet dans la page de discussion sur « la méthode d'écriture » et dans la page de la politique officielle à ce sujet. Seront commentées dans cette partie les recommandations qui figurent sur la page officielle Ṭarīqa al-kitāba طريقة الكتابة 'méthode d'écriture' (Wikipédia, 2025c) avec d'éventuelles références à des éléments étant apparus dans la discussion à ce sujet sur la page Niqāsh wīkībīdyā : ṭarīqat al-kitāba نقاش ويكيبيديا:طريقة الكتابة 'discussion Wikipédia : méthode d'écriture', laquelle a été archivée en trois parties auxquelles il sera fait référence par 'archive 1' (Wikipédia, 2020a), 'archive 2' (Wikipédia, 2020b) et 'archive 3' (Wikipédia, 2020c) respectivement, avec les noms d'utilisateur des contributeurs, quand ceux-ci ont signé leur intervention.

En haut de la page officielle Ṭarīqa al-kitāba طريقة الكتابة 'méthode d'écriture' figure la recommandation : « l'égyptien peut s'écrire de n'importe quelle manière tant qu'un Égyptien peut le comprendre. » (Ghaly, le 23/07/2008) Pourtant, Vb4ever exprime son désaccord dans la discussion le 16/05/2010 (archive 3) :

ومش عجبي حكاية اكتب الحرف الي يعجبك زي ما واحد قال برضوا. لازم نحدد طريقة كتابة معينة وقواعد معينة. لازم. لازم كلنا نتناقش في الخجات ديه ونعمل حاجة ستاندارد للغة.

Ça ne me plaît pas cette histoire d'écrire la lettre qui te plaît, comme quelqu'un d'autre l'a dit. Il faut déterminer une méthode d'écriture spécifique et des règles spécifiques. C'est nécessaire. Nous devrions tous discuter de ces questions et créer un standard pour la langue.

Sur la page officielle, la problématique de la multiplicité des manières d'écrire apparaît également comme assez importante pour mériter cinq injonctions sur les treize phrases du paragraphe intitulé نصائح 'conseils', dont :

لو سمحت ماتغيرش في طريقة الاملا في المقالات. طول ما مقاله كلها مكتوبه بطريقة املا واحده. ولو عاوز تزود عليها. اكتب بنفس طريقة الاملا بتاعتها.

S'il te plaît, **ne modifie pas le type d'orthographe** des articles, tant que l'article entier est rédigé dans un seul type d'orthographe. Si tu souhaites y ajouter quelque chose, écris dans le même type d'orthographe. [mise en gras d'origine] (Esperfulmo, le 10/02/2010)

Une phrase liste les graphes utilisables dans les titres, à savoir :

أ إ آ ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن ه و ي ء ي
چ et ف پ
Sont aussi autorisés dans le corps des articles les graphes

Dans le paragraphe suivant, intitulé تنظيم 'organisation', une section intitulée الحروف 'les lettres' liste les instructions graphe par graphe.

Les remarques concernant ces graphes seront étudiées dans les parties "recommandations officielles" traitées dans les choix orthographiques faits sur Wikipédia maşri (voir chapitre 4). Pour chacun, la règle et sa traduction seront indiquées, suivie des éventuels commentaires et références à la discussion.

Après la section intitulée الحروف 'les lettres', figurent de petites sections dans le paragraphe intitulé تنظيم 'organisation' dont la suivante :

«Le point et la virgule : le point et la virgule sont très importants. Fais-y attention pour que personne ne se perde au milieu d'une conversation. Ils s'écrivent collés au mot qui a été écrit avant eux, et suivis d'une espace.» (Ghaly, le 04/09/2008)

3.4 Automatisation des choix orthographiques

Comme dans les autres éditions, plusieurs contributeurs de Wikipédia maşri ont créé des programmes, certains appelés "bots", définis comme "des agents automatiques ou semi-automatiques qui interagissent avec Wikipédia comme le fait un utilisateur, mais pour des tâches répétitives et fastidieuses pour un humain" (Wikipédia, 2025g). Certains appliquent des modifications "cosmétiques" (modifiant le texte du code sans modifier le texte visible sur la page lue), d'autres interviennent sur le texte de l'article même, parfois avec un effet normatif sur l'écriture.

3.4.1 Ghalybot

En janvier 2009, Ghaly, le Wikipédien à l'origine du projet Wikipédia maşri, a commencé à utiliser son "bot", nommé GhalyBot, essentiellement pour des modifications techniques (liens avec des articles dans d'autres langues) (Wikipédia, 2024b). Plus récemment, il a assigné à Ghalybot des tâches de "correction" du texte des articles. Au 25 août 2024, le programme, publié sur la page autowikiGhalyBot, comporte près de 4.800 instructions visant à trouver (*find*) une suite de caractères et à les remplacer (*replace*) par une autre suite (Wikipédia, 2025e). Certaines visent à remplacer des suites de caractères latins (noms propres) par une translittération en caractères arabe. D'autres visent à "égyptianiser" (tamşīr - تمصير) le texte écrit, notamment des expressions perçues comme relevant de la fuṣḥā. Ce "bot" a un impact considérable puisque en avril 2025, le nombre des ses contributions s'élève à plus de 4,8 millions, soit plus de 40% de l'ensemble des contributions faites sur Wikipédia maşri, ce qui signifie que chaque article de l'édition a été modifié au moins une fois par Ghalybot (Wikipédia, 2025f).

Ghaly semble aussi utiliser son "bot" dans sa stratégie pour améliorer la faible performance de Wikipédia maşri pour ce qui est du critère de profondeur d'édition des articles. Beaucoup de ses récentes contributions visent à développer le texte des nombreux articles "ébauches" que contient cette édition de l'encyclopédie en ligne. Pour ce faire, il crée ou développe des articles en traduisant leur équivalent en anglais vers de l'arabe fuṣḥā, opération qu'il commente par la phrase "... أنشئت بترجمة الصفحة "...". "Créé en traduisant la page "...". Pour traduire, il semble utiliser la solution mise en place récemment sur Wikipédia proposant un outil d'aide à la traduction des articles d'une édition vers une autre édition, ou alors simplement un traducteur en ligne. L'anglais est choisi comme langue source probablement parce que les articles sont souvent déjà bien développés et c'est une langue que

maîtrise Ghaly. L'arabe fuṣḥā est choisi comme langue cible vraisemblablement de par sa disponibilité parmi celles du traducteur automatique et sa proximité avec "la langue égyptienne moderne", malgré les dénégations des créateurs de Wikipédia maṣri, dont Ghaly, à ce sujet. Après publication de cette version en arabe fuṣḥā dans l'édition maṣri, Ghalybot intervient pour "égyptianiser" le texte. Un exemple de ce type d'opérations, daté du 8 décembre 2023, 12:04, est présenté dans le tableau 3.6 pour le début de l'article *Deafness* 'Surdité', intitulé طرش /tʕaraʃ/ sur Wikipédia maṣri. La version après "égyptianisation" par Ghalybot est celle du 23 février 2025, 14:48, après 11 interventions de ce même "bot" et 2 interventions d'un autre "bot", ces dernières n'ayant pas affecté le texte présenté.

Édition en anglais	
Deafness has varying definitions in cultural and medical contexts. In medical contexts, the meaning of deafness is hearing loss that precludes a person from understanding spoken language, an audiological condition.[1] In this context it is written with a lower case <i>d</i>. It later came to be used in a cultural context to refer to those who primarily communicate through sign language regardless of hearing ability, often capitalized as Deaf and referred to as "big D Deaf" in speech and sign.[2][3] The two definitions overlap but are not identical, as hearing loss includes cases that are not severe enough to impact spoken language comprehension, while cultural Deafness includes hearing people who use sign language, such as children of deaf adults.	
Édition Wikipédia maṣri	
Traduction vers arabe fuṣḥā	Après "égyptianisation" par Ghalybot
للص-مم تعريفات مختلفة في السياقات الثقافية والطبية. في السياقات الطبية، معنى الطرش هو فقدان السمع الذي يمنع الشخص من فهم اللغة المنطوقة، هيا حالة سمعية . في السياق هذا يتم كتابته بالحرف الصغير d . تم استخدامه لاحقاً في سياق ثقافي للإشارة إلى أولئك الذين يتواصلون بشكل أساسي من خلال لغة الإشارة بغض النظر عن القدرة على السمع ، وفي الغالب ما يتم كتابتهم بالطرش "د الطرش الكبار" د الطرش "في الكلام والإشارة. يتداخل التعريفان لكن غير متدورين، حيث يشمل فقدان السمع الحالات التي مش شديدة بما يكفي للتأثير على فهم اللغة المنطوقة، في حين يشمل الطرش الثقافي سماع الأشخاص الذين يستخدمون لغة الإشارة، مثل أطفال البالغين الطرش .	للص-مم تعريفات مختلفة في السياقات الثقافية والطبية. في السياقات الطبية، معنى الطرش هو فقدان السمع الى يمنع الشخص من فهم اللغة المنطوقة، هيا حالة سمعية . في السياق ده يتم كتابته بالحرف الصغير d . تم استخدامه بعدين في سياق ثقافي للإشارة لدول الى يتواصلون بشكل أساسي منفى لغة الإشارة بغض النظر عن القدرة على السمع ، و في الغالب ما يتم كتابتهم بالطرش ويشار ليهم باسم "الطرش الكبار" د الطرش " في الكلام والإشارة. يتداخل التعريفان لكن غير متدورين، حيث يشمل فقدان السمع الحالات الى مش شديدة بما يكفي للتأثير على فهم اللغة المنطوقة، في حين يشمل الطرش الثقافي سماع الأشخاص الى يستخدمون لغة الإشارة، زى أطفال البالغين الطرش .

Tab. 3.6: Développement d'un article de l'anglais vers la fuṣḥā puis la 'āmmiyya

Le remplacement le plus fréquent est la séquence ي suivi d'une espace (c'est-à-dire le

yā' final, non directement suivi d'une ponctuation) qui devient ى suivi d'une espace, soit la 'alif maqṣūra, considérée comme la manière égyptienne d'écrire cette finale (في devient (الثقافي devient الثقافى, يكتفى devient يكتفي).

Le pronom personnel هي devient هيا (mais هو n'est pas modifié), les relatifs الذي, الذين et التي deviennent الى, le démonstratif هذا devient ده (qu'une absence d'espace ultérieure colle au mot يتم suivant), la préposition خلال devient فى (qu'une absence d'espace antérieure colle au mot من précédant), ... Certains de ces changements permettent effectivement d'introduire des éléments d'arabe égyptien de façon pertinente, notamment le relatif الى. Pour d'autres les choix semblent plus aléatoires. Ainsi, l'expression من خلال, 'à travers, par' est transformé en من فى (si on fait abstraction de l'absence d'espace), et le sens voulu est perdu. Aussi problématique est la transformation de متطابقين 'identiques' (*identical* en anglais) en متدورين, un barbarisme qui s'explique par la transformation voulue de la fuṣḥā طابق en 'āmīyya دور 'étage', sans prendre la précaution de considérer un contexte d'emploi différent. Certaines des erreurs, notamment celles concernant les espaces, ont depuis été corrigées dans le code de Ghalybot, mais celles ayant été publiées subsistent. Le texte final obtenu est donc un mélange entre un fond fuṣḥā (notamment l'auxiliaire passif يتم, تم, le passif يشار, le pluriel du verbe يستخدمون, le duel sujet التعريفان, ...) et des ajouts 'āmīyya, parfois artificiels. Il ne s'agit plus ici de code-switching volontaire, de code-mixing involontaire, ni même de la "fuṣḥāmīyya" décrite par Rosenbaum (2000), avec une intention stylistique. Le texte produit en l'occurrence est une tentative d'écriture à grande échelle d'arabe égyptien, qui consiste en un compromis entre une base fuṣḥā¹¹ et un saupoudrage d'éléments 'āmīyya, sans travail d'édition affiné ultérieur.

L'impact majeur de Ghalybot sur les articles de Wikipédia maṣri justifie que le traitement de la langue opéré par ce programme soit décrit en détail lors de la présentation des choix orthographiques faits sur Wikipédia maṣri (voir chapitre 4, les sections "Traitement opéré par Ghalybot"). Son étude permet en effet de mettre en évidence les conceptions du contributeur majeur qu'est Ghaly en ce qui concerne la manière dont "la langue égyptienne moderne" doit être écrite. Il applique donc massivement certains principes prescriptifs et vise ainsi à établir une sorte de norme d'écriture de la langue utilisée sur Wikipédia maṣri.

3.4.2 Autres programmes

Parmi les autres programmes visant à modifier la manière d'écrire dans les articles de Wikipédia maṣri, il y a aussi celui de HitomiAkane (publié sur la page arzAutocorrect.js), qu'il décrit comme un script d'autocorrection spécifique pour la langue égyptienne, qui corrige certaines lettres lors de l'écriture pour être cohérent avec la méthode d'écriture en langue égyptienne. Ce programme se concentre essentiellement sur la transformation du yā' final ي en 'alif maqṣūra ى ou de la tā' marbūṭa ة en hā' ه (Wikipédia, 2020d). Son action n'étant pas enregistrée comme une contribution à un article, il est difficile d'évaluer son impact, mais son créateur est très actif sur Wikipédia maṣri avec plus de 1,8 million de

¹¹Cette base fuṣḥā, elle-même issue d'une traduction automatique de l'anglais, est loin d'être authentique, le traitement du sigle 'd' en anglais devenant 'd' puis د, par exemple, sans parler de la suite pratiquement intraduisible 'often capitalized as Deaf and referred to as "big D Deaf"' qui devient la suite obscure وفي الغالب ما يتم كتابتهم بالطرش وشار لهم باسم "الطرش الكبار" د الطرش.

contributions, soit près de 18% du total, ce qui fait de lui le premier contributeur humain (après deux “bots”) en avril 2025, de par sa technique de création automatisée d’articles en masse (cf. p. 48).

Le contributeur Redaking a aussi créé son “bot” Redakingbot, actif depuis juillet 2008 et auteur de plus de 180.000 contributions. Certaines de ces contributions modifient le texte des articles et aboutissent à des choix orthographiques dont certains seront discutés ultérieurement.

3.5 Conclusion

Le projet Wikipédia maṣri a désormais atteint une certaine maturité, avec un grand nombre d’articles à son actif, même s’ils ont tendance à manquer de profondeur. Ghaly, son initiateur, est toujours un de ses contributeurs les plus actifs, et son “bot” imprime sa marque sur l’ensemble des articles. Le renfort plus récent de HitomiAkane a permis de démultiplier le nombre d’articles, qu’il s’agit maintenant d’approfondir. Cette encyclopédie en ligne représente tout de même un corpus très précieux, composé de nombreux textes écrits en arabe égyptien de manière revendiquée par des rédacteurs multiples. C’est un véritable recueil d’innovations grapholinguistiques visant, par l’emploi du système d’écriture arabe, à transcrire la phonologie de l’arabe égyptien, mais aussi à affirmer sa différence par rapport à la norme fuṣḥā.

Chapitre 4

Choix orthographiques sur Wikipédia maşri

Ce chapitre vise à étudier la réalité de l'écriture de l'arabe égyptien par les rédacteurs des articles de Wikipédia en arabe égyptien, au regard de la situation phonologique et graphémique de la langue, ainsi que des recommandations faites sur cette édition de l'encyclopédie en ligne.

Outre l'observation de stratégies d'écriture particulières lors de la lecture de certains articles, une tentative de relevés statistiques de certains choix orthographiques a été réalisée par la méthode suivante : le fichier contenant la version de chaque article de Wikipédia maşri au 01/02/2024 a été téléchargé à partir de la page <https://dumps.wikimedia.org> pour obtenir arwiki-20240201-pages-articles.xml (3,5 Go). Dans le logiciel 010 Editor, des recherches de chaînes de caractères ont été faites pour obtenir à chaque fois un nombre d'occurrences pour chacune des graphies possibles d'un mot. Les résultats sont indiqués dans les tableaux de ce chapitre, en nombres (#) et en pourcentages (%).

4.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya

Au vu des recommandations et des arbitrages d'écriture faits par les "bots", il a semblé intéressant d'observer la fréquence d'utilisation des graphes sur Wikipédia maşri (arz) et de la comparer à celle de Wikipédia al-'arabiyya (ar). Parallèlement au fichier arwiki-20240201-pages-articles.xml (3,5 Go), le fichier arwiki-20240720-pages-articles.xml (11 Go) contenant la version au 20/07/2024 des articles de Wikipédia al-'arabiyya a été téléchargé. Une recherche de chacun des graphes a été faite dans chacune des éditions. Les graphes sont classés du moins représenté au plus représenté sur arz par rapport à ar dans les tableaux en annexe 4.1 et 4.2 respectivement. Les résultats seront étudiés en détail au sein de la plupart des sections suivantes dans la partie 'Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya'.

graphe	Wikipédia maşrı (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ا	17.263	0,0022%	43	4.027.633	0,1433%	35	233,31
آ	611	0,0001%	47	129.168	0,0046%	44	211,40
آ	1.575	0,0002%	46	235.193	0,0084%	43	149,33
آ	15.490	0,0020%	44	2.063.197	0,0734%	40	133,20
آ	4.596	0,0006%	45	451.382	0,0161%	42	98,21
ا	348.229	0,0444%	36	19.730.202	0,7018%	25	56,66
ذ	169.397	0,0216%	38	9.507.807	0,3382%	31	56,13
ر	53.702	0,0069%	40	2.616.609	0,0931%	37	48,72
م	51.656	0,0066%	41	2.382.296	0,0847%	39	46,12
م	43.806	0,0056%	42	1.159.890	0,0413%	41	26,48
آ	2.275.450	0,2904%	28	54.292.275	1,9312%	16	23,86
ؤ	152.032	0,0194%	39	2.868.430	0,1020%	36	18,87
ة	5.646.607	0,7206%	23	100.960.306	3,5911%	10	17,88
ث	1.008.905	0,1288%	31	12.334.500	0,4387%	30	12,23
آ	221.304	0,0282%	37	2.488.371	0,0885%	38	11,24
ئ	1.573.673	0,2008%	29	94.74.864	0,3370%	32	6,02
خ	5.017.158	0,6403%	25	28.551.540	1,0156%	23	5,69
ء	1.268.363	0,1619%	30	6.591.327	0,2345%	33	5,20
ش	7.213.905	0,9206%	22	34.717.311	1,2349%	22	4,81
غ	3.226.804	0,4118%	27	15.515.668	0,5519%	27	4,81
ح	9.334.034	1,912%	21	44.379.696	1,5786%	19	4,75
ظ	972.875	0,1242%	32	4.427.141	0,1575%	34	4,55
ل	63.034.942	8,0443%	3	275.542.547	9,8009%	2	4,37
ق	12.080.564	1,5417%	19	52.283.490	1,8597%	17	4,33
ف	20.309.171	2,5918%	14	87.907.105	3,1268%	11	4,33
ض	3.299.101	0,4210%	26	13.122.675	0,4668%	28	3,98
د	23.302.609	2,9738%	12	86.999.405	3,0945%	12	3,73
ب	28.206.045	3,5995%	10	102.078.112	3,6309%	9	3,62
Corpus	783.599.634	100%	47	2.811.387.118	100%	47	3,59

Tab. 4.1: Sous-représentation des graphes arz par rapport à ar

graphie	Wikipédia maşri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
Corpus	783.599.634	100%	47	2.811.387.118	100%	47	3,59
ي	71.586.498	9,1356%	2	255.872.222	9,1013%	3	3,57
و	46.553.694	5,9410%	6	161.904.072	5,7589%	5	3,48
ر	45.472.582	5,8030%	7	156.650.474	5,572%	6	3,44
ت	41.468.036	5,2920%	8	141.732.511	5,0414%	8	3,42
ع	23.291.652	2,9724%	13	79.194.474	2,8169%	14	3,4
ن	47.052.142	6,0046%	5	155.964.726	5,5476%	7	3,31
ا	118.929.723	15,1774%	1	391.184.188	13,9143%	1	3,29
ز	5.613.067	0,7163%	24	18.326.115	0,6519%	26	3,26
س	27.247.443	3,4772%	11	85.551.920	3,0431%	13	3,14
م	55.993.480	7,1457%	4	174.238.132	6,1976%	4	3,11
ط	9.366.993	1,1954%	20	25.908.792	0,9216%	24	2,77
ك	20.292.778	2,5897%	15	55.471.216	1,9731%	15	2,73
ص	16.558.167	2,1131%	16	45.069.246	1,6031%	18	2,72
ح	15.821.354	2,0191%	18	36.108.823	1,2844%	21	2,28
ه	31.729.113	4,0491%	9	38.498.437	1,3694%	20	1,21
ى	16.225.512	2,0706%	17	12.808.495	0,4556%	29	0,79
ل	493.200	0,0629%	34	29.921	0,0011%	45	0,06
ف	598.199	0,0763%	33	21.100	0,0008%	46	0,04
ج	456.134	0,0582%	35	14.114	0,0005%	47	0,03

Tab. 4.2: Sur-représentation des graphes arz par rapport à ar

4.2 Voyelles courtes

4.2.1 État des lieux

4.2.1.1 Situation phonologique

Aux trois voyelles /i/, /u/ et /a/ de la fuṣḥā, sont souvent ajoutés les phones alternatifs respectifs [e], [o] et [ɑ] dans la description de l'arabe égyptien.

4.2.1.1.1 Voyelles fermées /i/ [e] et /u/ [o]

Parallèlement à /i/ et /u/, les voyelles [e] et [o] sont citées dans certaines grammaires ou dictionnaires qui tendent à leur attribuer un statut de phonèmes, et fournissent quelques paires minimales où elles apporteraient une distinction déterminante : /betna/ 'notre maison' contre /bitna/ 'Nous avons passé la nuit' et aussi /ʔotˤtˤi/ 'ma chambre' contre /ʔutˤtˤi/ 'mon chat' (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 316). Voir aussi Girod et Aziz (2016, p. 10-11). Badawi et Hinds (1986, p. xvii-xviii) précisent que les voyelles courtes /e/ et /o/ résultent généralement de la contraction des voyelles longues respectives /e:/ et /o:/ dans une syllabe fermée : /ʕeʃna/ 'notre pain' < /ʕe:/ʃ/ 'pain' et /dorna/ 'notre tour' < /do:r/ 'tour'¹. Watson (2002, p. 22) indique cependant que ces deux voyelles courtes ne sont pas attestées chez la majorité des locuteurs du parler du Caire, et elle ne les considère pas comme des phonèmes distincts.

4.2.1.1.2 Voyelles ouvertes /a/ [ɑ]

Badawi et Hinds (1986, p. xvii) utilisent le symbole [ɑ] pour distinguer la voyelle ouverte postérieure non arrondie de sa jumelle antérieure, symbolisée par /a/. (Voir aussi Stephens et Maurice (2004, p. xi) et Youssef (2013, 97 sqq.) entre autres). Ils fournissent la paire minimale barr 'prendre soin de' - barr 'être fidèle (à une promesse)'. Watson (2002, p. 21) préfère pour ce cas parler d'une consonne battue alvéolaire pharyngalisée (emphatique) qu'elle transcrit ʀ et décrit donc la paire minimale ainsi : barr 'prendre soin de' - barr ʀ 'être fidèle (à une promesse)'. Elle n'envisage donc qu'un seul phonème voyelle /a/ (court ou long) et plutôt une série de phonèmes « marginaux » dont les consonnes pharyngalisées (emphatiques) des mots suivants (notées par un point sous le symbole) : ḥāḥa 'pape, pontife, patriarche' - bāba 'deuxième mois du calendrier copte', mayyiti 'mon eau' - mayyiti 'mon mort'. Il est à noter que toutes ces consonnes que Watson considère comme emphatisées apparaissent au voisinage d'une ou de deux voyelles /a/. Badawi et al. préfèrent plutôt différencier celles-ci dans leur transcription : bāba (48)-bāba (49), mayyit (842)-mayyit (839) et ʿalla (33)-walla (956). La relative rareté de ces paires minimales et l'hésitation à choisir quel élément du contexte, consonne ou voyelle, doit être érigé en phonème, incite donc à la prudence avant de franchir un tel pas.

L'existence en arabe égyptien de phonèmes voyelles courtes autres que ceux de la fuṣḥā ne fait donc pas l'objet d'un consensus.

¹En plus des contractions du type de celles décrites ci-dessus, Farnawany (1981, p. 32-33) cite /emta/ 'quand' et /ojma/ 'sculpture, bas-relief' et semble donc considérer /e/ et /o/ comme des phonèmes, là où Badawi et Hinds (1986, 35 et 46) écrivent /ʔimta/ et /ʔujma/.

4.2.1.1.3 Voyelles raccourcies

Certaines voyelles courtes sont issues du raccourcissement de voyelles longues. Ainsi, alors qu'en fuṣḥā le terme مصابيح 'lampes' se prononce /masʕa:bi:h/, en ʿāmmiyya il se prononce /masʕabi:h/, avec raccourcissement de la première voyelle longue pour respecter la règle qui veut qu'on ne conserve qu'une voyelle longue dans un mot (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 317).

Le raccourcissement de voyelles longues a lieu dans d'autres situations. En effet, en ʿāmmiyya, une voyelle longue ne peut pas être immédiatement suivie de deux consonnes (ibid., 324 sqq.), ce qui peut arriver par l'ajout d'un suffixe comme dans /be:t/ 'une maison' > /betna/ 'notre maison'. De même, l'ajout d'un suffixe avec une voyelle longue raccourcit celle déjà en place : /be:t/ 'une maison' > /bete:n/ 'deux maisons'. Finalement, une voyelle longue doit être accentuée, ce qui n'est parfois plus possible lorsque l'ajout d'un suffixe déplace l'accent : /be:dʕa/ 'un œuf' > /be'diṭha/ 'son œuf (à elle)'. Après avoir indiqué tous ces exemples, et bien d'autres, Abdel-Massih, Abdel Malek et Badawi (ibid., p. 327) précisent : « Puisqu'il est complètement prédictible, le raccourcissement discuté ci-dessus n'a pas à être indiqué dans l'écriture. »².

Parallèlement, les voyelles longues finales de la fuṣḥā sont généralement raccourcies en ʿāmmiyya. C'est notamment le cas des particules négatives /ma/ et /la/, de la préposition /ʕala/, des pronoms personnels suffixes de première personne pluriel /na/ et troisième personne singulier /ha/. Ceci inclut les terminaisons fuṣḥā des pluriels du type /a:ʔ/ dans /zumala:ʔ/ qui devient en ʿāmmiyya /zumala/ 'collègues' (Badawi et Hinds, 1986, p. 379). Ceci concerne aussi les verbes défectueux, ceux dont la radicale finale est yā' ي ou wāw و correspondant à des voyelles longues en fuṣḥā, qui se transforment en voyelles courtes en ʿāmmiyya : fuṣḥā : /rama:/ - /jirmi:/ > ʿāmmiyya : /rama/ - /jarmi/ 'jeter' (ibid., p. 353), fuṣḥā : /radʕa:/ - /jardʕu:/ > ʿāmmiyya : /raga/ - /jargu/ 'espérer' (ibid., p. 329).

4.2.1.1.4 Voyelles épenthétiques

En arabe égyptien, une voyelle épenthétique est automatiquement insérée après la deuxième de trois consonnes consécutives pour briser cette suite. Cette voyelle est analysée comme étant /i/, plus courte et plus relâchée selon Abdel-Massih, Abdel Malek et Badawi (2009, p. 322) qui donnent les exemples /ʕufti ʕa:gil/ 'j'ai vu un homme', /ma-kaltif/ 'je n'ai pas mangé'. Elle peut être accentuée : /kal'bina/ 'notre chien'. Ils précisent que cette voyelle épenthétique devient /u/ devant les pronoms suffixes -hum et -kum : /kal'buhum/ 'leur chien', /kal'bukum/ 'votre chien', et /a/ devant le pronom suffixe -ha : /kal'baha/. Ils notent : « Sa représentation graphique étant totalement redondante, la voyelle épenthétique n'est, en règle générale, pas indiquée par l'écriture. »³. Harrell (1957, p. 60) pointe la différence entre /kunti fe:n/ 'tu (féminin) étais où ?' et /kunte fe:n/ 'tu (masculin) étais où ?', ainsi qu'entre /ʔummu kwayyisa/ 'sa mère est bonne' et /ʔummo kwayyisa/ 'une bonne mère', d'où il déduit deux voyelles courtes épenthétiques /e/ et /o/, en parallèle des trois voyelles courtes phonémiques /i/, /a/ et /u/.

²"Since it is completely predictable, the shortening discussed above does not have to be indicated in the script."

³"Since its graphic representation is totally redundant, the helping vowel is not, as a rule, indicated by the script."

4.2.1.1.5 /a/ ou /(i)t/ final issu de la tā' marbūṭa

La tā' marbūṭa (tā' liée), ة, est une marque, toujours en finale d'un mot, qui peut notamment indiquer, en fuṣḥā, le féminin : طبيب /tʕabi:b/ 'un docteur' > طبيبة /tʕabi:ba/ 'une doctoresse', ou le nom d'unité à partir d'un nom générique : سمك /samak/ 'du poisson' > : سمكة /samaka/ 'un poisson'.

En 'āmmiyya, l'utilisation est la même, et la réalisation à l'oral est /a/, comme ci-dessus à la pause en fuṣḥā, ou /(i)t/ en premier élément d'une annexion : /tʕabi:bit binti/ 'la doctoresse de ma fille', avec élision du /i/ si cela ne conduit pas à une suite de plus de deux consonnes : /tʕabibt ibni/⁴ 'la doctoresse de mon fils'.

On trouve la même prononciation avec les pronoms clitiques à sens de possessifs : /tʕabibitha/ 'sa doctoresse (à elle)' et /tʕabibtu/ 'sa doctoresse (à lui)'. Cependant, la tā' marbūṭa peut aussi s'appliquer à un participe actif dont le pronom clitique aura sens de complément objet. Ainsi, on trouve : /fa:him/ 'comprenant (masculin)' > /fahma/ 'comprenant (féminin)' > /fahma:ha/ 'la comprenant (féminin)' (Aldrich, 2012, p. 109). La réalisation orale de la tā' marbūṭa est donc ici un /a:/ long, et non plus /(i)t/, malgré l'ajout d'un pronom suffixe.

Ces diverses réalisations orales de la terminaison tā' marbūṭa pourront donner lieu à des traitements différenciés à l'écrit en arabe égyptien.

4.2.1.1.6 Tanwīn

Le système des suffixes de déclinaison, qui servent à marquer la désinence des noms et adjectifs indéterminés, est quasiment inopérant dans les dialectes. Ainsi, les désinences /in/ et /un/ ne sont pas prononcées en arabe égyptien. Seule la fathātān, la désinence /an/, subsiste pour certains adverbes comme /ʔabadan/ 'jamais', /dajman/ 'toujours' ou /taʔri:-ban/ 'environ', /tʕabʕan/ 'bien sûr' sans véritable rôle syntaxique cependant. À noter que les locutions fuṣḥā /ʕasʕban ʕan/ et /naḍran ʕalajja/ deviennent en arabe égyptien /ʕasʕbin ʕan/ 'en dépit de' (Badawi et Hinds, 1986, p. 624) et /nadrin ʕalajja/ 'je jure (de faire telle ou telle chose)' (ibid., p. 855).

4.2.1.2 Réalisation graphique

Comme indiqué par Farnawany (1981, p. 50) et Rosenbaum (2004, 294 sqq.), les véritables voyelles courtes, celles non issues du raccourcissement d'une voyelle longue, ne sont généralement pas écrites en arabe égyptien, sauf sporadiquement pour éclairer la prononciation d'un mot, comme en arabe littéral. Sont alors utilisés les signes diacritiques de la fuṣḥā, à savoir la kasra pour /i/, la ḍamma pour /u/ ou et la fatḥa pour /a/ ou /a/.

Parfois, des voyelles courtes non issues du raccourcissement d'une voyelle longue sont tout de même transcrites par les graphes utilisés en fuṣḥā pour les voyelles longues, à savoir yā ي pour /i:/, wāw و pour /u:/ et 'alif ا pour /a:/. Rosenbaum (ibid., p. 297-299) donne notamment les exemples suivants : سَيِّي au lieu de سَيِّ/sitti/ 'madame', شَوَفَت au lieu de شَوَفَت /ʃuft/ 'j'ai vu' et ما حَدَش au lieu de ما حَدَش /ma ḥaddiʃ/ 'personne'.

⁴Le /i:/ long est raccourci en /i/ car il est ici suivi de deux consonnes, pour éviter la formation d'une syllabe extra-longue en milieu de mot (Gadalla, 2000, p. 34-36).

Par ailleurs, morphologiquement, des phénomènes ont aussi cours en arabe égyptien qui éloignent la prononciation de la transcription des mots équivalents en arabe littéral. Ainsi, certaines voyelles courtes sont issues du raccourcissement de voyelles longues, notamment par l'ajout d'un suffixe. Farnawany (1981, p. 50-51) donne les exemples de prononciations et graphies suivantes :

- /a:/ /xa:l/ خال 'oncle' > /a/ /xalna/ خالنا 'notre oncle'
- /i:/ /taʕli:m/ تعليم 'éducation' > /i/ /taʕlimhum/ تعليمهم 'leur éducation'
- /e:/ /be:t/ بيت 'maison' > /e/ /betkum/ بيتكم 'votre maison'
- /u:/ /ʔu:l/ قول 'dis' > /u/ /ʔulna/ قولنا 'nous avons dit'⁵
- /o:/ /lo:n/ لون 'couleur' > /o/ /lonha/ لونها 'leur couleur'

Il s'agit ici de garder une cohérence graphique entre le mot sans suffixe et celui avec suffixe, et donc de privilégier la transcription de la morphologie au détriment de la phonologie. L'absence des graphes yā ي, wāw و et 'alif ا rendrait la lecture plus difficile, ne serait-ce que parce que ce n'est pas la pratique à laquelle les lecteurs sont habitués, notamment en arabe littéral.

Il est à noter que dans les exemples donnés, les graphes yā ي, wāw و et 'alif ا sont employés en 'āmmiyya pour transcrire des voyelles courtes qui portent l'accent, qu'elles soient véritables ou issues de raccourcissement. Sachant que les voyelles longues 'āmmiyya sont décrites comme systématiquement accentuées (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 317), il est permis de se demander si ce phénomène ne reflète pas une convergence entre voyelles courtes accentuées et voyelles longues pour ne former qu'un seul phonème.

En position finale, les voyelles courtes sont issues de diverses situations, ce qui explique la multiplicité de leurs graphies. La plus simple est l'absence de signe graphique. Ainsi, Farnawany (1981, p. 58) cite انت /ʔinta/ 'tu (masculin)'⁶. L'emploi de diacritiques est aussi attesté par Rosenbaum (2004, p. 299) qui cite انت /ʔinti/ 'tu (féminin)' marqué par kasra. Le /i/ final correspond ainsi souvent à la marque du féminin, qui peut aussi être transcrit par yā ي comme dans انتي /ʔinti/ 'tu (féminin)' cité par Rosenbaum (ibid., p. 298) ou par 'alif maqṣūra اى, c'est-à-dire un yā final sans ses deux points, comme l'écrit Farnawany (1981, p. 138) : انتى /ʔinti/ 'tu (féminin)'.

La 'alif support de la faṭḥatān est généralement écrite, le plus souvent sans la faṭḥatān, comme indiqué par Rosenbaum (2004, p. 299) avec دايما /dajman/ 'toujours' ou par Farnawany (1981, p. 64) avec طبعاً /tʕabʕan/ 'bien sûr'. Les alternatives spécifiques à l'arabe égyptien que sont /yasʕbin ʕan/ 'en dépit de' et /nadrin ʕalajja/ 'je jure (de faire telle ou telle chose)' sont respectivement transcrites par غصباً ou غصب, غصب ou غصبين et ندراً ou ندرن selon Rosenbaum (2004, p. 299).

La voyelle /u/ finale se trouve dans la terminaison du pluriel des verbes, comme dans بعثوا /baʕatu/ 'ils ont envoyé' (Farnawany, 1981, p. 132), où elle est marquée par la suite

⁵Cet exemple s'apparente à شوفت de Rosenbaum (2004, p. 298) et n'est donc pas vraiment un raccourcissement d'une voyelle longue, sauf à considérer que l'arabe fuṣṣā فُصِّلَ lui-même est aussi issu d'un raccourcissement.

⁶La forme انت peut aussi correspondre à /ʔinti/ 'tu (féminin)' comme c'est le cas en fuṣṣā.

wa wāw-'alif fāriqa, comme en fuṣḥā, ce dernier graphe pouvant aussi être omis comme dans قالو لي /ʔalu:li/ 'ils m'ont dit' (Farnawany, 1981, p. 110). La voyelle /u/ est aussi la prononciation du pronom personnel ou du possessif masculin singulier suffixé, comme dans ارضه /ʔardʕu/ 'sa terre' (ibid., p. 132), où elle est marquée par le graphe ه, comme en fuṣḥā, où celui-ci correspond à la prononciation /hu/ ou /hi/. Dans ce même cas, elle peut aussi être marquée par le graphe wāw و, comme dans سيبو /si:bu/ 'laisse-le', voire par analogie avec le suffixe verbal, avec la suite وا wāw-'alif fāriqa comme dans روحلوا /ruḥlu/ 'va à lui' (ibid., p. 132), voir aussi Rosenbaum (2004, p. 298).

Pour /a/ final, outre l'absence de signe graphique, on peut concevoir le recours au diacritique fatḥa comme dans انت /ʔinta/ 'tu (masculin)' pour désambiguïser le genre, comme en fuṣḥā. Par ailleurs, le /a/ final est communément marqué par la tā' marbūṭa ه, qui correspond souvent au féminin, comme dans مدينة /madi:na/ 'ville' cité par Farnawany (1981, p. 127), qui explique que les deux points peuvent être omis pour donner ه, soit le graphe utilisé en fuṣḥā pour transcrire /h/, comme dans معلمه /miʕallima/ 'maîtresse d'école' (ibid., p. 127). La fusion des graphes ة et ه rend la graphie suivante ambiguë : مدرسه = son enseignant /mudarrisu/ ou une enseignante /mudarrisa/. Par analogie, cette graphie est utilisée sans lien avec une tā' marbūṭa fuṣḥā, comme dans عينيه /ʕenajja/ 'mes (deux) yeux' (Rosenbaum, 2004, p. 297), créant une homographie avec ʕene:(h)/ 'ses (deux) yeux', apparenté à ʕajnajhi/ en fuṣḥā.

Finale, /a/ final peut être transcrit par la 'alif ا, comme dans بعنا /biʕna/ 'nous avons vendu', ou par la 'alif maqṣūra ي, comme dans على /ʕala/ 'sur', parfois écrit علا avec la 'alif traditionnelle, phénomène que Farnawany (1981, p. 57) explique ainsi : "La majorité des écrivains suivent dans ce cas le système arabe standard, à savoir l'orthographe avec - ي -, tandis qu'une minorité essaie intentionnellement une nouvelle orthographe. Certains d'entre eux souhaitent par exemple que chaque /a/ à la fin d'un mot soit indiqué par - ا -".⁷

L'utilisation de plusieurs graphes pour transcrire une même voyelle et d'un même graphe pour diverses voyelles, notamment celles en fin de mot, illustre la tension entre le respect de l'étymologie ou de la morphologie fuṣḥā et le désir de représenter la phonologie 'āmmiyya, voire de revendiquer un droit à la différenciation.

4.2.2 Orientations Wikipédia maṣri

4.2.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle Ṭarīqa al-kitāba طريقة الكتابة 'méthode d'écriture' (Wikipédia, 2025c) mentionnent les points suivants :

4.2.2.1.1 Diacritiques

Les signes diacritiques servent notamment à transcrire les voyelles courtes, mais uniquement quand cela est jugé nécessaire :

⁷"Die Mehrheit der Schriftsteller folgt in diesem Fall dem hocharabischen System, nämlich der Schreibung mit - ي -, während eine Minderheit absichtlich eine neue Schreibweise versucht. Manche von ihnen wollen, z.B., daß jedes /a/ am Ende eines Wortes durch - ا - bezeichnet werden soll."

علامات التشكيل

علامات التشكيل ساعات بتكون مهمه لو نفس الكلمه كان مش واضح نطقها من السياق و ليها اكثر من معنى, لو انتطقت بكرا طريقه.

”Les signes diacritiques sont parfois importants si la prononciation d’un même mot n’est pas claire d’après le contexte et s’il a plus d’un sens, s’il se prononce de plusieurs manières.” (Esperfulmo, le 02/04/2009).

Les recommandations concernant les diacritiques restent ainsi très succinctes. Ils ne sont cependant pas écartés, l’objectif de leur utilisation étant la désambiguïsation d’éventuels homographes.

4.2.2.1.2 ‘alif maqṣūra

Pour la ‘alif maqṣūra ي, la recommandation est la suivante :

ي لما مش بيكون في حالة الألف اللينه, حرف ال"ي" هو الشكل التقليدي في اخر الكلام و الى مانغيرش في مصر. كُتابته من غير نقطتين اخر الكلام هي الطريقه الى سكان وادي النيل متعودين عليها. "ي" او "ي" هما حرف واحد. المتماشى عليه في ويكيبيديا مصرى هو كُتابته من غير نقطتين في اخر الكلام.

Lorsque ce n’est pas un cas de ‘alif layyina, la lettre ي est la forme traditionnelle à la fin des mots et qui n’a pas changé en Egypte. L’écrire sans deux points à la fin est la méthode à laquelle les habitants de la vallée du Nil sont habitués. ي ou ي sont une seule lettre.

Ce qui est cohérent dans Wikipédia égyptien, c’est de l’écrire sans deux-points à la fin du mot.

"احنا في : justification suivante est donnée :
"Nous, à l’école, on nous a appris à écrire yā’ [final] sans deux-points, et tous les journaux et publications égyptiens l’écrivent sans deux-points.” et il ajoute un lien vers une reproduction du billet de 1 livre égyptienne, où figure la mention البنك المركزي المصري /il-bank ul-markaziyy ul-mis’riyy/ ‘la banque centrale égyptienne’ (Ramsis II, 01/08/2008)⁸. Ce qui est perçu

⁸Cette tendance égyptienne à écrire le yā’ final sans points est illustrée dans un article de l’auteur égyptien "المشكلة الكبيرة - بالنسبة لي - بدأت عندما بدأت في الكتابة ، فكلما أرسلت مقالاً ، أو قصة ، تحتوي كلماتها على : ‘Ādil ‘Attiā : ياء أخيرة بدون نقطتين ، فإن الناشر غير المصري ، يقوم بوضع النقطتين ! وإذا وضعت النقطتين فإن الناشر المصري يقوم بحذفها !" "Le gros problème - pour moi - a commencé lorsque j’ai commencé à écrire. Chaque fois que j’envoyais un article ou une histoire comportant un « yā’ » final sans ses deux points, l’éditeur non égyptien les mettait ! Et si je mettais les points, l’éditeur égyptien les supprimait !". (Attiā, 2010). La suppression des deux points sous le yā’ final, fusionnant sa représentation avec la ‘alif maqṣūra, s’oppose pourtant à la décision de l’Académie de la langue arabe du Caire prise en 1980 dans laquelle elle préconise : "ترسم الألف اللينة بصورة الياء (غير منطوقة)", "أما الياء فتتقط للفرق. 'L’alif maqṣūra est dessinée comme un yā’ (non pointé), tandis que le yā’ est pointé pour [faire] la différence." (Amīn et Tarzi, 1984, p. 313). On observe cependant que les éditeurs de cette publication de l’Académie ne respectent pas leur propre règle, car aucun yā’ final n’apparaît pointé dans la page où elle est écrite, y compris pour la suite de mots الفارسي وهذا هو رأي أبي على الفارسي (et cela est l’avis de ‘Ābī ‘Alī l-Fārisī) (un célèbre grammairien de l’école d’al-Baṣrah). Dans *A Dictionary of Egyptian Arabic : Arabic-English*, tous les yā’

comme la tradition et l'identité égyptienne semble justifier le choix de la fusion de deux graphes différents aux réalisations phonologiques distinctes, au risque de créer des ambiguïtés, comme le signale le même Ramsis II le 7/12/2008 (archive 2) en pointant la différence entre علي /ʕali/ ‘‘Ali’’ (le nom propre) et على /ʕala/ ‘sur’ (la préposition).

4.2.2.1.3 tā’ marbūṭa

Pour la tā’ marbūṭa ة (appelée ‘‘tā’ de féminisation’’), la recommandation est la suivante :

ته التائيث

ته التائيث بتكتب حسب النطق. في حالة الاضافه لما بتتنطق زى ال"ت": ة/؛ او /ه/ في الحالات الثانيه.

«La tā’ de féminisation s’écrit selon sa prononciation. En état d’annexion lorsqu’elle se prononce comme le ت : [elle s’écrit] ة ; ou ه dans les autres cas.»

Cette règle est mentionnée par Lagrange (2013, p. 750) qui la classe parmi les ‘‘usages anciennement installés de l’écriture du dialecte’’, tout en notant qu’elle n’est appliquée que de manière aléatoire dans le roman de Muṣṭafā Muṣarrafa *Qanṭara al-laḏī Kafar* (1965) ‘Qantara qui a perdu la foi’. Par ailleurs, dans la page de discussion (archive 1), Ramsis II propose le 02/12/2008 pour ‘la superficie de l’Égypte’ la graphie المساحه بتاعة مصر /il-misa:ħa bita:ʕit masˤr/, hors état d’annexion pour المساحه /il-misa:ħa/, et مساحة مصر /mi-sa:ħit masˤr/ en état d’annexion, mais il suggère aussi مساحت, pour rendre la prononciation précise de /misa:ħit masˤr/. Mahmudmasri suggère le 07/12/2008 (archive 1) de ne pas utiliser la tā’ marbūṭa à la place de ه pour le son /a/ final du démonstratif masculin singulier /da/ : ده et non دة ‘ça’ et كده et non كدة /kida/ ‘comme ça’.

4.2.2.2 Traitement opéré par Ghalybot

4.2.2.2.1 Diacritiques

Ghalybot (cf. p. 57) cherche à faire disparaître les diacritiques. Il vise notamment la fathātān et sa ’alif finale, au moins quand la première précède la seconde, pour tous les graphes portant cette fathātān, sauf pour دَا، ذَا، شَا et غَا qui ne sont pas modifiés, mais بَا, par exemple, est transformé en ب. Quand la fathātān est sur la ’alif, il remplace par la consonne correspondante seulement عَا، فَا، قَا، لَا، مَا et نَا qui sont donc remplacés dans les deux cas. Il est difficile de comprendre la raison de cette sélection. Cela permet à un adverbe courant comme أبداً ou أبداً ‘jamais’ de ne pas être remplacé.

Toujours est-il que ces effacements de fathātān et ’alif finaux ne sont pas sans poser des problèmes de compréhensibilité. Ainsi le titre du film égyptien دائما معك ‘Toujours avec toi’ tel qu’il apparaît sur son affiche de 1954 a été automatiquement transformé en دايم معك

finaux sont écrits sans points, même s’ils se prononcent /i/ comme pour نسي /nisi/ ‘oublier’ (Badawi et Hinds, 1986, p. 861). Ce phénomène s’observe également en fuṣḥā, par exemple dans *A Dictionary of Modern Written Arabic* dans lequel le verbe correspondant /nasija/ et son maṣdar /nasj/ sont transcrits نسي (Wehr et Cowan, 1979, p. 1130).

sur la page de l'actrice **كامل عايد** 'Aida Kamel' après que le rédacteur Raafat l'eut écrit avec **fatḥatān** **عايداً** معاك. L'adverbe 'toujours' est ainsi devenu l'adjectif 'permanent' et le sens de l'expression est brouillé.

Ghalybot cherche aussi à transformer une trentaine de passifs vocalisés **fuṣḥā** en passifs dialectaux avec préfixe **ات**. Il remplace donc **قُتِلَ** par **اتقتل** "il a été tué" et **وُلِدَ** par **اتولد** "il est né", effaçant en passant les diacritiques. De même, il élimine des **kasratān** en remplaçant l'expression **لحد ما** "dans une certaine mesure" par **شويه** "un peu" et la suite **لـ** par **لى**, pour transformer, par exemple, **ليال** "nuits" en **ليالى**, la forme 'āmmiyya.

Finalement, Ghalybot, vise l'expression **ما يُفسّر** "ce qui s'explique" et la transforme en **الى يتفسر** /ʔilli jitfassar/, éliminant les diacritiques **ḍamma**, **fatḥa** et **šadda**.

4.2.2.2.2 'alif maqṣūra

Ghalybot remplace le **yā** **ي** suivi d'une espace, qui transcrit traditionnellement /i/ en fin de mot, par une 'alif maqṣūra **ى**, ceci pour refléter la tradition d'écriture égyptienne, telle que décrite sur la page officielle **Ṭarīqa al-kitāba** **طريقة الكتابة** 'méthode d'écriture' (cf. page 69). Il s'agit d'une opération de remplacement massive, qui apparaît sur chaque article de l'encyclopédie, au risque de perdre de manière irréversible la distinction entre **yā** final et 'alif maqṣūra.

4.2.2.2.3 tā' marbūṭa

D'une part, Galybot supprime une ou plusieurs **tā' marbūṭa** dans environ 180 cas, notamment en remplaçant **لفترة وجيزة** par **لمدة صغيرة** 'pour une courte durée', et aussi **الثورة الفرنسية** par **الثوره الفرنسيه** 'la révolution française'. Il applique la règle selon laquelle la **tā' marbūṭa** hors d'état d'annexion, se prononçant /a/, s'écrit sans ses deux points. D'autre part, Galybot ajoute une ou plusieurs **tā' marbūṭa** dans environ 370 cas, notamment en remplaçant **عام في** par **سنة** 'en l'an', mais aussi **مملكة البرتغال** par **مملكة پورتوجال** /mamlakit bur-tuga:l/ 'le royaume du Portugal', appliquant ainsi la règle qui veut que la **tā' marbūṭa** en annexion, et donc prononcée /it/, s'écrit avec ses deux points.

4.2.3 Observations dans le corpus

4.2.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya

4.2.3.1.1 Diacritiques

L'ensemble des diacritiques est bien moins utilisé sur Wikipédia maṣri que sur Wikipédia al-'arabiyya, comme l'illustre le tableau 4.3.

Parmi les graphes les moins fréquents sur Wikipédia maṣri on trouve le **sukūn** **◌ْ**, qui sert à marquer l'absence de voyelle entre deux consonnes. Sa contribution à désambiguïser des homographes est probablement ressentie comme assez basse.

La **fatḥatān**, ou désinence /an/, est le diacritique du **tanwīn** le plus utilisé sur Wikipédia maṣri, mais il est le moins fréquent par rapport à Wikipédia al-'arabiyya. Cette rareté

graphie	Wikipédia maṣri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
◌◌̣	17.263	0,0022%	43	4.027.633	0,1433%	35	233,31
◌◌̣̣	611	0,0001%	47	129.168	0,0046%	44	211,40
◌◌̣̣̣	1.575	0,0002%	46	235.193	0,0084%	43	149,33
◌◌̣̣̣̣	15.490	0,0020%	44	2.063.197	0,0734%	40	133,20
◌◌̣̣̣̣̣	4.596	0,0006%	45	451.382	0,0161%	42	98,21
◌◌̣̣̣̣̣̣	53.702	0,0069%	40	2.616.609	0,0931%	37	48,72
◌◌̣̣̣̣̣̣̣	51.656	0,0066%	41	2.382.296	0,0847%	39	46,12
◌◌̣̣̣̣̣̣̣̣	43.806	0,0056%	42	1.159.890	0,0413%	41	26,48
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.3: Fréquence des diacritiques sur arz et ar

relative résulte probablement de l'intervention de Ghalybot (cf. p. 57) visant à le faire disparaître.

Dans Wikipédia maṣri, les deux autres diacritiques du tanwīn /in/ et /un/ sont très rares et apparaissent essentiellement dans des citations du Coran, de la Bible ou d'autres textes sacrés. Les trois diacritiques des voyelles courtes َ, ُ et ِ restent relativement moins rares sur arz, notamment pour désambiguïser la prononciation d'un mot, comme le veut aussi l'usage en fuṣḥā.

4.2.3.1.2 'alif maqṣūra

Pour la 'alif maqṣūra, les occurrences sont tout simplement plus nombreuses sur Wikipédia maṣri que sur Wikipédia al-‘arabiyya, malgré un ratio général de 3,59 en faveur de cette dernière, comme illustré dans le tableau 4.4.

graphie	Wikipédia maṣri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ي	71.586.498	9,1356%	2	255.872.222	9,1013%	3	3,57
ى	16.225.512	2,0706%	17	12.808.495	0,4556%	29	0,79
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.4: Fréquence de ي et ى sur arz et ar

Cela reflète la pratique d'écriture attribuée à la tradition égyptienne qui consiste à écrire tout yā' ي final sans ses deux points, donnant ainsi lieu à un graphe unique ى qui outre sa valeur de 'alif maqṣūra prononcée /a/ acquiert celle de yā' finale prononcée /i/ ou /j/ en arabe égyptien. L'influence de l'opération quasi-systématique de remplacement effectuée par Ghalybot est ici clairement évidente. Pour autant, le graphe ي reste aussi, voire plus, fréquent sur Wikipédia maṣri, et ne souffre pas de son élimination de la finale des mots.

4.2.3.1.3 tā' marbūṭa

Le tableau 4.5 révèle que parmi les graphes relativement bien plus présents sur Wikipédia maṣri, on trouve ٥.

graphe	Wikipédia maṣri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
٥	5.646.607	0,7206%	23	100.960.306	3,5911%	10	17,88
ه	31.729.113	4,0491%	9	38.498.437	1,3694%	20	1,21
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.5: Fréquence de ٥ et ه sur arz et ar

En l’occurrence, ce graphe ne correspond pas seulement à la lettre hā’ de la fuṣḥā, mais aussi à la tā’ marbūṭa ٥, que les recommandations de Wikipédia maṣri enjoignent d’écrire sans ses deux points lorsqu’elle n’est pas à la fin d’un mot en état d’annexion, et donc qu’elle se prononce /a/. En toute logique, la tā’ marbūṭa ٥ est bien moins utilisée sur Wikipédia maṣri.

4.2.3.2 Exemples de choix graphiques

Du point de vue graphique, l’usage bien établi en fuṣḥā de ne pas noter les voyelles courtes, sauf exceptions, notamment afin de désambiguïser un terme, est largement repris par la tradition d’écriture en arabe égyptien. De plus les trois diacritiques kasra , ḍamma et fatha , utilisés pour transcrire respectivement les trois voyelles /i/, /u/ et /a/ de la fuṣḥā, s’avèrent insuffisants pour transcrire les phones alternatifs respectifs /e/, /o/ et /a/ en ‘ām-miyya.

4.2.3.2.1 Traitement des phones alternatifs

4.2.3.2.1.1 /a/ ou /a/ Dans Wikipédia maṣri, la discrimination entre les voyelles ouvertes /a/ antérieure et /a/ postérieure (courtes ou longues) n’est pas marquée à l’écrit, ni celle entre les consonnes emphatiques ”marginales” et leurs équivalents traditionnels non emphatiques en fuṣḥā. Les quelques cas de paires minimales éventuelles ne sont pas homographes pour autant, des éléments du contexte, notamment la terminaison, permettant de distinguer les mots. Ainsi, /majjit/ ‘mort’ est écrit سيرة الظاهر بيبرس ميت (Roman de Baïbars)⁹ et /majja/ ou /mʕajja/ مياه (Eau). De même, /ba:ba/ ‘deuxième mois du calendrier copte’ est écrit باب (بابه) et /ba:ba/ ou /bʕa:bʕa/ ‘pape’ est écrit بابا (بابا Pape). Aucune recommandation pour marquer cette différence n’a été trouvée dans les pages de discussion sur la méthode d’écriture. Le diacritique suggéré par Ḥalīl ‘Asākir pour transcrire cette voyelle postérieure (cf. tableau 1.1) n’est pas mentionné (Asākir, 1955, p. 185). Son absence parmi les caractères des claviers arabes, ainsi que celle d’une alternative à la ‘alif explique probablement le renoncement à cette discrimination.

⁹Le titre des articles d’où les exemples sont tirés sera indiqué entre parenthèses, avec une éventuelle traduction.

4.2.3.2.1.2 [e] et [o] Pour [e] et [o], leur classification en tant que phonèmes distincts en arabe égyptien ne fait pas consensus. Ces voyelles courtes ne seront donc transcrites dans cette étude que lorsqu'elles seront issues de la voyelle longue correspondante, à savoir pour [e] la voyelle /e:/ longue : /be:t/ 'maison' donnant /betha/ 'sa maison' (Badawi et Hinds, 1986, p. 114) et pour [o] la voyelle /o:/ longue : /no:ma/ 'lieu de couchage' donnant /nomtak/ 'ton lieu de couchage' (ibid., p. 893). Dans le corpus, ces voyelles sont habituellement transcrites respectivement par yā' et wāw, les graphes traditionnellement utilisés pour les voyelles longues d'origine. Ainsi, on trouve بيتها /betha/ (كلوديا كاردينالي) Claudia Cardinale) et نومتها /nomtaha/ (روميو و جوليتت) Roméo et Juliette).

4.2.3.2.2 Position initiale

En début de mot, les voyelles s'appuient généralement sur le "coup de glotte" /ʔ/, marqué par la 'alif accompagnée, ou non, de la hamza. Cette situation sera décrite dans la section 4.4.

4.2.3.2.3 Position médiane

4.2.3.2.3.1 /i/ Pour /i/, Voyelle antérieure non-arrondie (semi-)fermée, on trouve de manière exceptionnelle le diacritique kasra utilisé pour préciser par exemple la prononciation de فضل /fidʕil/ 'il continua' (صلاح ذو الفقار Salah Zoufcar) pour le distinguer de فضل /fadʕil/ 'grâce' (نهج البرده), ou la prononciation de حجر /higr/ 'genoux, giron' (لاب توب) Ordinateur portable) pour le distinguer de حجر /hagar/ 'pierre' (عبر المعونه Eben Ezer), par exemple. Le graphe yā' ي, malgré sa valeur traditionnelle de transcription de la voyelle longue, subsiste quand celle-ci est raccourcie en arabe égyptien, par exemple par l'ajout d'un suffixe formant une suite de deux consonnes : سيبيه /si:bu/ 'laisse-le' (مغور) > sibhim /sibhum/ 'laisse-les' (نشيد الاستقلال التركي İstiklâl Marşı). On trouve aussi cette lettre utilisée dans les verbes concaves de voyelle /i/ aux personnes du passé autres que les troisièmes, comme جيت /gibt/ 'j'ai apporté' (3 occurrences dans (ليا نظرة), probablement par réminiscence du présent avec la voyelle longue يجيب /bijgi:b/ 'il apporte'. La graphie جبت reste cependant plus courante avec au moins 14 occurrences dans le corpus, dont محمد الموجي) Mohammed al-Mougui).

4.2.3.2.3.2 /u/ Pour /u/, voyelle postérieure arrondie fermée, on trouve de manière exceptionnelle le diacritique damma utilisé pour préciser par exemple la prononciation de فريق المغرب لكورة /ʔusu:d/ 'lions' (أُسود) et de قدامك /ʔudda:mak/ 'devant toi' (الربيع اغنيه)) (Équipe du Maroc de football). Le graphe wāw و, malgré sa valeur traditionnelle de transcription de la voyelle longue, subsiste quand celle-ci est raccourcie en arabe égyptien par l'ajout d'un suffixe déplaçant l'accent : فوق /fo:ʔ/ 'dessus' > فوقانيه /foʔa'nijja/ 'du dessus, supérieur' (هليكوبتر Hélicoptère). On trouve aussi cette lettre utilisée dans les verbes concaves de voyelle /u/ aux personnes du passé autres que les troisièmes, comme قلت /ʔult/ 'j'ai dit' (29 occurrences dont une dans (حلو يا حلو), probablement par réminiscence du présent avec la voyelle longue يقول /bijʔu:l/ 'il dit'. La graphie قلت reste cependant

plus courante avec au moins 238 occurrences dans le corpus (dont une dans سَمْسِمِيَّة (آلة) Simsimiyya).

De manière moins justifiée, à côté de la transcription de la négation courante مَش /miʃ/ ou /muʃ/ '(ne) pas'¹⁰, on trouve parfois la forme مَوْش (dont une dans رِيَال يَمْنِي Riyal yéménite, parmi 626 occurrences (dont de nombreuses transcriptions de noms étrangers), mais jamais مِيش. Selon Farnawany (1981, p. 131), avec مَوْش, il s'agit de préciser la prononciation d'un mot qui peut se prononcer soit avec /i/ soit avec /u/.

4.2.3.2.3.3 /a/ Pour /a/, voyelle ouverte, on trouve des occurrences du diacritique fatḥa , notamment dans نَفَس /nafas/ 'souffle' (فَاصِلُهُ مَنْقُوطُهُ) avec deux fatḥa successives, probablement pour distinguer ce mot de نَفْس /nafs/ 'esprit; même' (عِلْمُ نَفْس). On trouve aussi نَوْر /nawar/ 'gitans' (غُرْ مِصْر Doms en Égypte) pour marquer la différence avec نور /nu:t/ 'lumière' (النُّور Lumière). Cet usage du diacritique fatḥa reste cependant rare.

Le graphe 'alif ʾ, malgré sa valeur traditionnelle de transcription de la voyelle longue, subsiste quand celle-ci est raccourcie en arabe égyptien, par exemple quand un participe actif de première forme comme سَامِع /'sa:miʕ/ 'entendant, qui entend' (حَدَّ سَامِعٌ حَاجَهُ؟) 'Quelqu'un a entendu quelque chose ?') est mis au féminin سَامِعَةٌ ou سَامِعُهُ /'samʕa/ (سَامِعَاتُ) ou au pluriel سَامِعِينَ /sam'ʕi:n/ (المُسَحَرَاتُ).

Il a été trouvé une occurrence où la voyelle épenthétique /a/ accentuée a été transcrite par la 'alif : اسْمَاها /ʔis'maha/ 'son nom (à elle)' (شَرَكَةُ لِيُون رُولَان par Raafat), ce qui contribue à expliquer la conservation de la 'alif pour سَامِعُهُ /'samʕa/, mais pas pour سَامِعِينَ /sam'ʕi:n/. Il semble cependant qu'une voyelle courte accentuée soit plus facilement marquée par une 'alif que celles non accentuées.

4.2.3.2.4 Position finale

4.2.3.2.4.1 /i/ Le traitement du phonème /i/ en final de mot donne lieu à des solutions scripturales variées. Pour ce qui est du féminin singulier de la deuxième personne du pronom personnel isolé ou de l'équivalent du verbe "être" au passé, comme le montre le tableau 4.6, sont hégémoniques les formes non marquées, c'est-à-dire celles qui ne permettent pas de différencier le masculin du féminin hors-contexte.

L'utilisation des diacritiques est anecdotique, au masculin ou au féminin. Un nombre plus élevé de recours à ي ou ى indique des tentatives de rendu phonographique de cette voyelle finale courte, par ces graphes, même s'ils transcrivent en fuṣḥā des voyelles longues. Ce dernier graphe ى reste ambigu, puisqu'il est utilisé ici comme alternative à ي, selon la pratique égyptienne déjà décrite (cf. 69, bien que sa valeur phonémique habituelle en fuṣḥā soit le /a:/ long. Théoriquement, il n'est pas exclu qu'il transcrive en fait le masculin /ʔinta/, même si tous les exemples consultés relevaient d'un contexte féminin, et donc d'une interprétation en /ʔinti/.

¹⁰/miʃ/ et /muʃ/ viennent respectivement des expressions arabes féminin مَا هِيَ شَيْءٌ /ma: hiʒa fajʔ/ et masculin مَا هُوَ شَيْءٌ /ma: huwa fajʔ/ 'ce n'est rien'. La distinction entre les deux genres a finalement été perdue, laissant /miʃ/ et /muʃ/ comme variantes interchangeables.

finale			‘étais’		‘tu’						Total	
		API	/kunt-/		/?int-/						#	%
s	∅	?	كُنْتُ	1003	أَنْتَ	1193	أَنْتِ	354	إِنْتَ	65	2615	90,5
sm		/a/	كُنْتُ	1	أَنْتَ	5	أَنْتِ	6	إِنْتَ	0	12	0,4
sf		/i/	كُنْتُ	3	أَنْتَ	3	أَنْتِ	10	إِنْتَ	1	17	0,6
	ي		كُنْتِي	5	أَنْتِي	44	أَنْتِي	16	إِنْتِي	11	76	2,6
	ى		كُنْتِي	6	أَنْتِي	145	أَنْتِي	7	إِنْتِي	12	170	5,9
	total		1018		1390		393		89		2890	100

Tab. 4.6: Graphies des finales de /kunt-/ et /?int-/

Il peut être intéressant de rapprocher ces résultats du traitement de la traduction habituelle de 'maintenant' en arabe égyptien, à savoir /dilwa?ti/ (Badawi et Hinds, 1986, p. 950) pour laquelle la répartition statistique des graphies se fait selon le nombre d'occurrences suivant : 90 pour دلوقت, 1 pour دلوقت, 126 pour دلوقتي et 3.727 pour دلوقتي¹¹. Le recours à ى est donc prédominant, renforcé par les transformations vers cette dernière graphie effectuées par Ghalybot des équivalents arabes حالياً et الآن.

Trois des principales prépositions se terminent par /i/, à savoir /fi/ 'dans' apparenté au littéral في /fi:/ avec /i:/ long et écrit séparé du mot suivant, /li/ 'à' et /bi/ 'par' dont les équivalents en littéral comportent la même voyelle courte (sauf /li/ qui devient /la/ devant un pronom clitique). Ces deux dernières prépositions sont cliticisées au mot suivant en fuṣḥā. Sur Wikipédia maṣri, comme observable dans le tableau 4.7, on trouve un nombre non négligeable de transcriptions de ces prépositions par le graphe consonantique seul, et donc dans sa forme isolée, précédée et suivie d'une espace. Une alternative (toujours entre deux espaces) est l'ajout après le graphe consonantique de la kaṣīda كشيده ou taṭwīl تطويل, ce trait horizontal utilisé ordinairement pour justifier un texte en allongeant les liaisons entre caractères arabes (Unicode U+0640). Son utilisation ici laisse entendre une liaison avec la suite, mais sans cliticisation, pour introduire une liste avec retour à la ligne, un mot avec lien hypertexte, dont les crochets du code html font obstacle à la cliticisation, etc. Le diacritique kasra n'a été observé qu'une fois dans ل pour toutes ces formes isolées précédentes. La multiplicité des contextes n'a pas permis de quantifier les cas de cliticisation, mais la lecture des articles permet de constater qu'ils sont courants, sauf pour /fi/ qui ne se cliticise pas non plus en littéral, et d'ailleurs Ghali déconseille la graphie فاوروبا /fi ? urubba/ 'en Europe' dans les recommandations d'écriture de Wikipédia maṣri (Wikipédia, 2025c).

Le deuxième type d'écriture de ces prépositions consiste à faire suivre le graphe consonantique par le graphe yā ي, ce qui correspond à la graphie fuṣḥā de /fi/. La transformation en 'alif maqṣūra ى, selon la tendance égyptienne, est même hégémonique dans ce cas. Pour /li/ et /bi/, cet ajout de yā ou alif maqṣūra correspond en littéral à l'ajout du pronom clitique de première personne du singulier pour exprimer 'à moi', 'par moi', etc. Cela peut être sa valeur aussi en arabe égyptien ou simplement la transcription de

¹¹ On trouve aussi 68 occurrences de دلوقتي correspondant à la prononciation /?ilwa?ti/. Ces termes sont dérivés de وقت /waqt/ 'temps' précédé de l'article défini et d'un préfixe démonstratif /di/ (Manfred Woidich, 2008, p. 14).

	‘dans’			‘à’			‘par’			total	
	/fi/			/li/			/bi/			#	%
		#	%		#	%		#	%		
Ø	ف	57.458	01%	ل	6.838	08%	ب	2593	09%	66.889	01
-	ف	45	00%	ل	63.163	82%	ب	6.724	25%	69.932	01
ي	في	21.166	00%	لي	792	01%	بي	346	01%	22.304	00
ى	في	4.203.782	98%	لى	5.577	07%	بى	17.143	63%	4.226.502	96
total		4.282.451	97%		76.370	01%		26.806	00%	4.385.627	100%

Tab. 4.7: /i/ final des prépositions

mots étrangers comme dans ‘BBC’ بي بي سي. Il n’a pas été trouvé d’utilisation en tant que prépositions seules de ses deux graphes consonantiques écrits avec yā’ ou alif maqṣūra de manière isolée.

L’écriture des prépositions /li/ et /bi/ suivies de pronoms clitiques se fait essentiellement de deux manières visibles dans le tableau 4.8. Ce peut être une juxtaposition, comme en arabe littéral, ce qui peut correspondre à la prononciation en arabe égyptien pour la troisième personne du masculin singulier /lu(h)/ et /bu(h)/ du féminin singulier /laha/ ou /liha/ et du pluriel /luhum/. On observe aussi l’insertion de yā’ ي, afin de refléter la prononciation en arabe égyptien par exemple pour la troisième personne du masculin singulier /li:(h)/ et /bi:(h)/, du féminin singulier /li:ha/ et /bi:ha/ et du pluriel /li:hum/ et /bi:hum/.

	ل			ب			total
	sm	sf	p	sm	sf	p	
Ø	له	لها	لهم	به	بها	بهم	
	4.533	687	869	33	17	186	6.325
	47%	8%	51%	17%	32%	43%	32%
ـي	ليه	ليها	ليهم	بيه	بيها	بيهم	
	4.939	7.070	828	157	36	245	13.275
	52%	91%	48%	82%	67%	56%	67%
total	9.472	7.757	1.697	190	53	431	19.600

Tab. 4.8: Prépositions /li/ et /bi/ avec pronoms clitiques de 3° personne

Les statistiques montrent que la deuxième solution, celle avec yā’ est préférée dans les deux tiers des cas. Cela reflète la prononciation la plus courante, et renforce la différenciation de l’écriture de l’arabe égyptien de celle de l’arabe littéral, avec la contribution de Ghalybot pour ليه, ليها et ليهم.

4.2.3.2.4.2 /u/ L’écriture de /u/ final est aussi soumise au choix entre morphologie et phonologie, essentiellement pour le traitement du pronom clitique masculin singulier qui en fuṣḥā est ه /hu/ et qui se transforme en /u/ en ‘ammiyya (cf. tableau 4.9).

Comme l’expose le tableau, la préférence pour l’écriture morphologique est forte, que

traduction	API	ه			و			total
			#	%		#	%	
son nom	/ismu/	اسمه	12.060	99	اسمو	49	01	12.109
son fils	/ibnu/	ابنه	1.379	99	ابنو	2	01	1.381
de lui	/minnu/	منه	2.083	97	منو	44	03	2.127
chez lui	/ʕandu/	عنده	1.299	96	عندو	49	04	1.348
comme lui	/zajju/	زيه	74	80	زيو	18	20	92
aussi	/bardu/	برده	52	13	بردو	324	87	376
	/bard ^h u/	برضه	669	82	برضو	146	18	815
total			17.616	96		632	04	18.248

Tab. 4.9: /u/ final transcrit par ه ou و

ce soit après un nom comme ابن اسم ou اسم, ou après une préposition. Avec /bardu/ ‘aussi’, qui vient du turc <bir de> (Badawi et Hinds, 1986, p. 66), la même hésitation apparaît, bien qu’ici aucune considération morphologique n’entre en compte. On observe ainsi que dans ce contexte, le graphe ه final acquiert la valeur phonémique /u/, alors que dans les cas où il est issu de la tā’ marbūṭa, sa valeur phonémique est /a/. Les choix orthographiques observés ne respectent donc pas l’isomorphisme, et seuls une minorité de rédacteurs choisit d’écrire phonographiquement le /u/ final avec wāw و.

Une différence entre fuṣḥā et ‘ammiyya est la prise en compte de al-’alif al-fāriqa placée après le wāw و final des formes plurielles du verbe à la 3e personne masculin de l’accompli فعلوا (cf. p. 38). Cette ’alif ne se prononce pas, et donc de nombreux rédacteurs ne l’écrivent pas sur Wikipédia maṣri, notamment pour /ka:nu/, mais elle reste majoritaire dans certains autres verbes, comme illustré dans le tableau 4.10.

traduction	API	وا			و			total
			#	%		#	%	
ils étaient	/ka:nu/	كانوا	798	18	كانو	3.507	82	4.305
ils firent	/ʕamalu/	عملوا	216	52	عملو	199	48	415
ils restèrent	/baɖu/	بقوا	177	60	بقو	117	40	294
ils allèrent	/ra:ħu/	راحوا	132	51	راحو	124	49	256
ils continuèrent	/fid ^h lu/	فضلوا	71	37	فضلو	116	63	187
ils s’assirent	/ʔaʕadu/	قعدوا	73	71	قعدو	29	29	102
ils envoyèrent	/baʕatu/	بعثوا	38	49	بعثو	39	51	77
total			1.505	26		4.131	74	5.636

Tab. 4.10: /u/ final transcrit par وا ou و

4.2.3.2.4.3 /a/ La transcription de /a/ final est déjà un enjeu en arabe littéral, puisque, outre le diacritique fatḥa َ, la tā’ marbūṭa ة se prononce ainsi à la pause (cf. p. 66). Le /a:/ long final peut également être transcrit de deux manières, à savoir par la ’alif traditionnelle ٱ (cf. p. 36) ou la ’alif maqṣūra ى (cf. p. 37). Le /a:/ ou /a:/ long fi-

traduction	API	ا			ه			ة			total
			#	%		#	%		#	%	
ce, ça	/da/	دا	9.668	27	ده	26.768	73	دة	146	00	36.582
comme ça	/kida/	كما	559	08	كده	6.825	90	كدّة	146	02	7.530
à l'intérieur (de)	/guwwa/	جوا	665	33	جوه	1.375	66	جوة	25	01	2.065
à l'extérieur (de)	/barra/	برا	1.764	58	بره	1.258	41	برة	20	01	3042
total			12.656	26		36.226	73		337	01	49.219

Tab. 4.11: /a/ final transcrit par ا, ه ou ة

nal devenant courts en arabe égyptien, c'est donc une multiplicité de choix qui s'offre au rédacteur wikipédien mašri, auxquels il faut ajouter la tendance, recommandée officiellement, de supprimer les deux points de la tā' marbūṭa quand elle se prononce effectivement /a/, c'est à dire en dehors d'une annexion, ce qui donne le graphe ه semblable à celui qui transcrit /h/ final en fuṣṣḥa et 'āmmiyya. La transcription de /a/ final est donc essentiellement réalisée par la 'alif traditionnelle, la 'alif maqṣūra, la tā' marbūṭa, avec ou sans ses deux points, ou par l'absence de graphe, étant donné qu'en 'āmmiyya le diacritique fatḥa est rarement marqué, et encore moins en fin de mot, puisqu'il n'y a pas de déclinaisons.

On observe ainsi une variété de choix graphiques pour certains mots courants dont la prononciation se termine par /a/.

Dans le tableau 4.11, les démonstratifs /da/ 'ce, ça', /kida/ 'comme ça' et les prépositions/adverbes /guwwa/ 'à l'intérieur (de)', /barra/ 'à l'extérieur (de)' peuvent être écrits avec la 'alif traditionnelle ou la tā' marbūṭa, avec ou sans ses deux points. Lors des recherches, une occurrence de /da/ démonstratif sans graphe après د a été trouvée dans le début de phrase ود ه يعني أن /wi da ha jaʕni ʔinni/ 'et ça signifiera que' (مثل مساعد Second rôle). Sinon, pour ces mots, sauf /barra/, le graphe privilégié pour rendre le /a/ final est hā' ه, ce qui correspond aux remplacements automatiques effectués par Ghalybot. Pour les démonstratifs, les résultats correspondent à l'étude diachronique de Avallone (2015), où elle note une disparition des formes avec la 'alif traditionnelle dans les extraits consultés les plus récents, même si cela diffère des mots apparentés en arabe littéral, à savoir هذا et كما. La graphie avec ه devient donc celle du masculin des démonstratifs en 'āmmiyya, alors que paradoxalement, en fuṣṣḥa, c'est la finale du féminin هذه 'cette' (qui devient en 'āmmiyya دي /di/). La graphie avec 'alif est cependant celle conservée en 'āmmiyya pour كذا /kaza/ 'plusieurs' (527 occurrences) en concurrence avec كرا (23) et كره (15) cette dernière forme essentiellement de la main d'un seul rédacteur : Masrialltheway.

Dans le tableau 4.12, on observe que pour les pronoms /huwwa/ 'il', /hijja/ 'elle' et /humma/ 'ils, elles', outre la 'alif traditionnelle ou la tā' marbūṭa, avec ou sans ses deux points, on trouve la graphie fuṣṣḥa sans graphe final (∅).

Dans les résultats de nos recherches, le graphe le plus répandu pour rendre le /a/ final est de loin la 'alif ا, là encore grâce au choix effectué par Ghalybot pour هيا. Bizarrement, il ne se préoccupe pas des autres pronoms. Par contre, dans ses articles créés en masse, HitomiAkane fait lui le choix de هوا. De son côté Redakingbot contribue à l'apparition d'un nombre non négligeable d'occurrences de هوه. L'utilisation de l'orthographe fuṣṣḥa sans graphe ajouté reste très courante en 'āmmiyya (moins pour هي dont le remplacement

traduction	finale	Ø		ا		ه		ة	
il	/huwwa/	هو	440.900	هوا	427.351	هوه	47.345	هوة	8
elle	/hijja/	هي	3.350	هيا	1.182.597	هيه	496	هية	9
		هى	105.891						
ils, elles	/humma/	هم	1.110	هما	4.351	همه	4	همّة	7
	total		551.251		1.614.299		47.845		24

Tab. 4.12: /a/ final transcrit par Ø, ا, ه ou ة

est fait par Ghalybot). Il n'y a en effet pas vraiment de logique à un ajout de graphe, puisque la différence phonologique d'avec la fuṣḥā est essentiellement le doublement de /w/ pour /huwwa/, de /j/ pour hijja/ et de /m/ pour /humma/ (avec l'ajout de /a/ court pour ce dernier). L'emploi de la šadda devrait suffire, mais elle reste marginale avec 10 occurrences seulement pour هو et aucune pour هي ou هم. Ces rédacteurs de Wikipédia en arabe égyptien s'éloignent donc de l'utilisation du script arabe en tant qu'abjad, et l'exploitent d'une manière phonographique plus poussée en explicitant la voyelle finale /a/. L'ajout de ces graphes finaux ressemble par ailleurs à un choix délibéré visant à marquer l'emploi de ces pronoms comme relevant du registre de la 'āmmiyya. On retrouve ici la dichotomie entre orthographe étymologique et phonémique mentionnée par Farnawany (1981, p. 4), opposition à laquelle fait aussi référence Rosenbaum (2004, p. 285). Le phénomène relève aussi de ce que Michalski (2019, p. 38-44) décrit en arabe marocain dans le cadre de la relation donneur-receveur¹² entre fuṣḥā et 'āmmiyya, cette dernière ayant emprunté son script à la première. Comme en arabe marocain, il est possible d'appliquer son principe de défiance envers le donneur¹³, car les choix d'écriture reviennent ici à distancier l'arabe égyptien de l'arabe littéral ne serait-ce que pour justifier l'existence même de l'édition maṣri.

Pour l'ensemble de ces mots, l'emploi de la tā' marbūṭa reste marginal, peut-être parce qu'on est loin de son rôle originel de marqueur de féminin ou de nom d'unité. De plus certaines de ces formes correspondent à des morphèmes différents, tels que /himma/ 'zèle' pour همّة.

Dans le tableau 4.13, on observe que pour les prépositions /šala/ et /hatta/ et le verbe /baʔa/ à l'accompli, /jibʔa/ à l'inaccompli, l'écriture fuṣḥā se fait avec la 'alif maqṣūra, en concurrence en arabe égyptien avec la 'alif traditionnelle et la tā' marbūṭa sans ses deux points.

Les occurrences avec la tā' marbūṭa et ses deux points de عة correspondent essentiellement à /šilla/ 'maladie, excuse', tandis que celles avec la tā' marbūṭa avec ou sans ses deux points حته et حتة correspondent à /hatta/ 'morceau, lieu'. Ces formes n'ont pas été prises en compte, seules celles ayant le même sens que celui exprimé par les autres graphies ont été comptées, en excluant les occurrences de بة /buʔʔu/ 'sa bouche'. Ici la graphie fuṣḥā reste majoritaire (hégémonique pour على. L'exception est با avec un recours à la 'alif traditionnelle majoritaire, qui ressemble à l'application du principe de défiance envers le donneur.

¹² "the donor-receiver relationship" (Michalski, 2019, p. 38) ou "donor-recipient" relationship (ibid., p. 26)

¹³ the principle of donor-defying (ibid., p. 43)

traduction	finale	ى	ا	ه	
sur	/ʕala/	على	علا	عله	7
jusqu'à	/ħatta/	حتى	حتا		
il est devenu	/baʔa/	بقى	بقا	بقه	1
il devient	/jibʔa/	يبقى	يبقا	يبقه	2
	total				
		296.502	4.413	10	

Tab. 4.13: /a/ final transcrit par ى, ا ou ه

Dans le monde arabe, une des manières de faire référence aux mois du calendrier grégorien est l'emprunt d'équivalents adaptés de langues européennes. Pour 'juin' et 'juillet' on trouve يونيو /junju/ et يوليو /julju/ 'juillet', deux termes qui servent de titre à leurs articles respectifs sur Wikipédia al-'arabiyya, dans lesquels il est précisé qu'on trouve en Égypte les formes alternatives يونيو /junja/ et يوليو /julja/ 'juillet'. Ce sont d'ailleurs ces dernières formes qui sont favorisées sur Wikipédia maṣri, notamment par les remplacements faits par Ghalybot, avec le nombre d'occurrences suivantes : 1.397 pour يونيو et 955 pour يوليو, contre 6.129 pour يونيو et 4.788 pour يوليو. Il semble que la perception de ces formes en /a/ soit associée à une certaine "égyptianité", et donc qu'elles soient favorisées.

La transcription de /a/ final est aussi un enjeu pour la marque du féminin notée par la tā' marbūṭa ة en fuṣḥā, notamment pour les formes de participe actif des verbes de forme 1, c'est-à-dire le schème /fa:ʕil/ - فاعل. En 'āmmiyya, ce schème masculin ne diffère pas mais, contrairement à la fuṣḥā, l'ajout au féminin de /a/ final transforme sa structure phonémique interne qui devient /faʕla/, /a:/ long devenant /a/ court accentué et /i/ disparaissant. Avec le mot très courant /ʕa:jiz/ signifiant 'il veut' et systématiquement écrit عايز (797 occurrences), la transcription du féminin /'ʕajza/ se fait essentiellement par l'ajout de la tā' marbūṭa ة عايزة (135) mais aussi de ه عايزه (207). Cette dernière orthographe est ambiguë car elle peut aussi correspondre au masculin suivi du pronom personnel masculin de 3^e personne, pour signifier 'il le veut' prononcé /'ʕajzu/. Les quelques occurrences consultées relevaient cependant du féminin, suivant la règle de l'effacement des deux points de la tā' marbūṭa. Une seule occurrence utilisant la 'alif : عايزا a été trouvée isolée, dans une page de discussion (ويكيبيديا:مقالات مختاره/ترشيحات/ثوره فرنسية). Une occurrence de la disparition de la 'alif de la première syllabe a aussi été trouvée pour le synonyme de /'ʕajza/, c'est-à-dire /'ʕawza/ orthographiée exceptionnellement عوزة (Pachnamounis). Il y a donc dans la grande majorité des cas conservation de l'orthographe morphologique avec 'alif pour la première syllabe et adaptation de la graphie fuṣḥā pour rendre compte de la phonologie pour la deuxième syllabe.

4.2.3.2.5 tā' marbūṭa et annexion

Il est intéressant d'observer le respect de la règle qui veut que la tā' marbūṭa s'écrive avec ses deux points quand elle transcrit la prononciation /(i)t/, et sans ses deux points quand elle transcrit la prononciation /a/.

Le tableau 4.14 montre que la règle est majoritairement respectée, d'abord avec les séquences remplacées par Ghalybot, comme /wila:ʒit ʒurʒja/ ولاية جورجيا 'L'État de Géorgie' et /il-ħarb il-ʕa:lamijja t-tanja/ الحرب العالميه الثانيه 'la Seconde Guerre mondiale'. Mais

traduction	API		#	%	correct ?	total
L'État de Géorgie	wila:ʃit ʒurʒja	ولاية جورجيا	859	62	oui	1.397
		ولاية جورجيا	538	38	non	
Le gouvernorat du Fayoum	muħafzʕit il-fajju:m	محافظة الفيوم	688	99	oui	696
		محافظة الفيوم	8	01	non	
Le Canal de Suez	qana:t is-siwe:s	قناة السويس	610	94	oui	652
		قناه السويس	42	06	non	
la Seconde Guerre mondiale	il-ħarb il-ʕa:lamijja t-tanja	الحرب العالمية الثانية	794	72	oui	1.117
		الحرب العالمية الثانية	323	28	non	
Le Caire	il-qa:hira	القاهرة	10.502	57	oui	18.608
		القاهرة	8.106	43	non	
Héliopolis	masʕr il-gidi:da	مصر الجديدة	326	60	oui	551
		مصر الجديدة	225	40	non	

Tab. 4.14: Fréquence de la tā' marbūṭa par rapport à hā'

la règle s'avère aussi majoritairement respectée dans les écrits des rédacteurs lambda.

On trouve aussi dans le corpus quatre occurrences de la tā' marbūṭa précédée du signe diacritique kasra pour expliciter sa prononciation /it/ comme pour شوية /ʃuwwajjit/ 'un peu de' (اللغة المصرية الحديثه) ou encore /muʔablit/ 'rencontre' اسبرانتو (Espéranto).

Quand la tā' marbūṭa est prononcée /it/, certains rédacteurs choisissent de la transformée en tā' simple ت. Ainsi, Adhm youssef écrit dans l'article ماكسيميان Maximien Hercule les expressions suivantes : ناحيت شمال افريقيا /naħjit ʃima:l ʔafriqja/ 'vers le nord de l'Afrique' (au lieu de ناحية), 285 سنت /sanit .../ 'l'année 285' (au lieu de سنة), 305 لغايت /liḡa:ʃit .../ 'jusqu'en 305' (au lieu de لغاية). WikiMichel écrit بتاعت الناس /luyɑ:t bita:ʕit in-na:s/ 'les langues des gens' (au lieu de بتاعة) (Langage de programmation) لغة برمجية) من Congo a pris fin' (au lieu de من ساعة ما خلصت حرب الكونغو Kinshasa).

Les chiffres de trois à neuf se prononcent avec la terminaison /t/, correspondant à la tā' marbūṭa, devant des noms pluriels commençant par une voyelle, y compris lorsque le "coup de glotte" est effacé comme pour /ʔajja:m/ 'jours', qui est perçu comme devenant /tijja:m/ (Stephens et Maurice, 2004, p. 32). Ainsi parallèlement aux graphies canoniques New York ou سبعة أيام Seven Days in Utopia), Samsam22 écrit /xamas tijja:m/ 'cinq jours' (شهر قبطيه) et /sabaʕ tijja:m/ 'sept jours' (سبع تيام). De même pour /ʔafhur/ 'mois (pluriel)', qui est perçu comme devenant /tuʃhur/ (ibid., p. 83). Samsam22 écrit /xamas tuʃhur/ 'cinq mois' (أشهر) (الأشرف طومان باي) al-Achraf Tuman Bay II).

De sont côté, Ramsis II a recours à la suite tā' + kaʕida ت placée entre deux espaces après le nombre et avant le nom compté comme dans /xamas tijja:m/ 'cinq jours' (تقويم بهائي).

Les graphies de la tā' marbūṭa en annexion visent donc souvent à refléter ses prononciations réelles en arabe égyptien, à travers de multiples stratégies.

4.2.3.2.6 Faṭḥatān

Pour ce qui est de la faṭḥatān, des exemples de répartitions de son utilisation apparaissent dans le tableau 4.15.

	'jamais'				'toujours'		'environ'		total	
API	/ʔabadan/				/dajman/		/taʔri:ban/		#	%
ا	أبداً	51	ابداً	68	دائماً	502	تقريباً	323	944	52
أ	أبداً	68	ابداً	30	دائماً	133	تقريباً	199	430	24
ل	أبداً	266	ابداً	11	دائماً	4	تقريباً	22	303	17
□	أبد	6	ابد	10	دائم	85	تقريب	32	133	7
	'éternité'				'permanent'		'approximation'		total	
	391				119		724		576	
									1810	
									100	

Tab. 4.15: Fréquence de la faṭḥatān

On constate que pour أبداً, l'emploi de la faṭḥatān reste très majoritaire, la suite de graphes دأ n'étant pas transformée en د par Ghalybot. Par contre, les suites ماً, ماً et باً sont remplacées par Ghalybot en م et ب respectivement, et donc favorisent les graphies دائماً et تقريب. Les graphies avec 'alif finale sans faṭḥatān semblent être le choix naturel d'une petite majorité de rédacteurs. Il faut noter que les formes sans 'alif ni faṭḥatān peuvent ainsi être des adverbes "corrigés" par Ghalybot, mais ont souvent leur propre sens.

On trouve des occurrences de عنييا بتضحك غصبا عني (Foie gras) et عنييا بتضحك غصبا عني (عينييا بتضحك), la prononciation /yas'ba/ 'malgré' étant possible à côté de /yas'bin/, qui est celle audible lors de l'écoute de cette dernière chanson. L'emploi de nūn est aussi attesté, par exemple dans عنييا بتضحك غصبا عني (Qait Bay), qui semble expliciter la prononciation /yas'bin/ en contradiction avec la règle fuṣḥā qui requiert la kasratān. On trouve ce diacritique dans la chanson مؤدب أصلي مؤدب غصبا عني /yas'bin ʕanni/ 'malgré moi' tel qu'écrit canoniquement par Raafat, et ensuite corrigé en عنييا بتضحك غصبا عني avec ajout de nūn pour expliciter la prononciation de manière redondante. Parallèlement, on trouve عنييا بتضحك غصبا عني 'je jure que' (سعاد) Souad Hosni) avec ici aussi explicitation de la prononciation /nadrin ʕalajja/ 'je jure que' par le nūn. L'utilisation de la kasratān est avérée dans la chanson عنييا بتضحك غصبا عني avec عنييا بتضحك tel qu'écrit par Raafat et ici encore corrigé par un anonyme avec ajout d'un nūn redondant pour عنييا بتضحك. Le désir de rendre visible cette prononciation égyptienne semble avoir particulièrement motivé ces rédacteurs.

4.2.3.2.7 Conclusion

Quoiqu'il en soit, la pratique généralisée d'écriture en abjad, c'est-à-dire sans marquer les voyelles courtes, évacue largement la question de l'écriture de celles-ci. Ce phénomène est renforcé dans Wikipédia maṣri par le fait que les rédacteurs y ont nettement moins recours aux diacritiques des voyelles courtes que ceux de Wikipédia al-'arabiyya. Cependant, transcrire la prononciation de la voyelle finale du mot apparaît souvent comme nécessaire, et de nombreuses solutions, souvent hétérogènes, sont adoptées. La diversité

des choix graphiques effectués pour transcrire cette voyelle finale montre cependant le peu d'importance accordé à l'élaboration de règles orthographiques concertées.

4.3 Voyelles longues

4.3.1 État des lieux

4.3.1.1 Situation phonologique

Phonologiquement, outre les trois voyelles longues de la fuṣṣḥā /a:/, /i:/ et /u:/, les descriptions de l'arabe égyptien s'accordent généralement pour faire des voyelles longues /e:/ et /o:/ des phonèmes à part entière (Spitta, 1880, p. 45) (Badawi et Hinds, 1986, p. XVII). On les trouve ainsi dans des mots d'origine étrangère : /karate:h/ 'karaté' (ibid., p. 742) d'origine japonaise, /radjate:r/ 'radiateur (automobile)' d'origine française (ibid., p. 333), /ʔo:dʕa/ 'salle, pièce' d'origine turque (ibid., p. 43) et /balko:n/ 'balcon' d'origine française (ibid., p. 100), etc. D'autre part, /e:/ et /o:/ se distinguent notamment dans les paires minimales : /di:n/ 'religion' - /de:n/ 'dette' et /mu:t/ 'meurs (impératif)' - /mo:t/ 'mort (concept)' (Harrell, 1957, p. 62). Comme le note Watson (2002, p. 22), /e:/ et /o:/ sont très souvent issues des «diphthongues»¹⁴ respectives /ay/ et /aw/ présentes en fuṣṣḥā : /dayn/ 'dette' et /mawt/ 'mort'. C'est le raccourcissement de ces voyelles longues /e:/ et /o:/ qui produit les voyelles courtes correspondantes /e/ et /o/ que l'on retrouve dans les paires minimales indiquées ci-dessus.

Il est à noter que les «diphthongues» /ay/ et /aw/ sont également attestées en arabe égyptien, ce qui donne lieu aux paires minimales suivantes : /ʃajla/ 'portant (féminin)' - /ʃe:la/ 'charge, fardeau' et /ʃawza/ 'voulant (féminin)' et /ʃo:za/ 'besoin, nécessité' (ibid., p. 23).

La différenciation à l'écrit de ces paires voyelles longues et «diphthongues» est donc un enjeu dans un système qui ne compte que deux graphes dédiés : yā' ي et wāw و.

Par ailleurs, on constate en dialecte un phénomène d'allongement des voyelles courtes. C'est ainsi le cas après les prépositions suivantes quand elles sont suivies d'un suffixe pronominal : /li/ 'à, pour' > /li:na/ 'à, pour nous', /bi/ 'avec' > /bi:k/ 'avec toi (masculin)' (Gadalla, 2000, p. 223-224).

Pour les verbes sourds, ceux dont la 2° et 3° radicales sont identiques, l'arabe dialectal ne suit pas le schéma régulier de l'arabe littéral pour la conjugaison de l'accompli aux premières et deuxième personnes, comme pour le verbe /ʃadd(a)/ 'tirer' : fuṣṣḥā : /ʃadadtu/, /ʃadadta/, /ʃadadti/, /ʃadadna:/, /ʃadadtum/ > 'āmmiyya : /ʃadde:t/, /ʃadde:t/, /ʃadde:ti/, /ʃadde:na/, /ʃadde:tu/. Il y a donc en arabe égyptien conservation de la gémination des 2° et 3° radicales et ajout avant les suffixes de conjugaison de la voyelle longue /e:/, similaire à celle de la conjugaison des verbes défectueux comme /rama/ > /rame:t/ (ibid., p. 66).

¹⁴Il semble qu'il s'agisse ici plutôt de la combinaison d'une voyelle et d'une glissante (*glide* en anglais), combinaison qui, à la différence d'une véritable diphthongue, n'est pas entièrement contenue dans le noyau de la syllabe, la glissante occupant la place de la coda. Le terme traditionnel «diphthongues» sera cependant utilisé ici, faute de mieux.

4.3.1.2 Réalisation graphique

La transcription des voyelles longues en arabe égyptien se fait essentiellement avec les graphes arabes traditionnels. Ainsi, /i:/ est transcrit par yā ي, que cette voyelle soit dans un mot arabe comme طويل /ṭawi:l/ 'long', ou dans un mot d'origine étrangère comme جيبه /ʒi:ba/ 'jupe' (Farnawany, 1981, p. 137), ou /e:/ comme dans بيت /be:t/ 'maison' ou بيه /be:(h)/ 'monsieur' (ibid., p. 141) (du titre turc ottoman *beg*). De même, /i:/ est transcrit par wāw و, que cette voyelle soit dans un mot arabe comme قوم /ʔu:m/ 'lève-toi' ou dans un mot d'origine étrangère comme بخور /bunʒu:r/ 'bonjour', ou /o:/ comme dans نوم /no:m/ 'sommeil' ou روب /ro:b/ 'peignoir' (ibid., p. 132-133).

/a:/ est transcrit par 'alif ا comme dans هات /ha:t/ 'apporte' (ibid., p. 56) ou جرنال /gur-na:l/ 'journal' (ibid., p. 121). Cette 'alif peut être surmontée de la madda آ pour transcrire la séquence /ʔa:/ comme dans آدم /ʔa:dam/ 'Adam' (ibid., p. 54). /a:/ peut ne pas être transcrit, comme en fuṣḥā, pour les mots comme الله /ʔalla:h/ 'Dieu' (Badawi et Hinds, 1986, p. 33) ou لكن /la:kin/ 'mais' (ibid., p. 776), même si Farnawany (1981, p. 54) indique la graphie phonologique alternative لاكن.

Par ailleurs, le graphe 'alif ا transcrit aussi la voyelle longue /a:/ du participe actif de première forme comme عايز /ʕa:jiz/ 'voulant, qui veut'. Quand ce participe actif est issu d'un verbe transitif, il peut régir un objet direct (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 14-15) (Mitchell, 1956, p. 101). Cet objet direct peut être réduit à un pronom personnel affixe هو عايزه /huwwa ʕajzu/ 'il le veut' (Aldrich, 2012, p. 95). La marque graphique du féminin de ce participe actif est la tā' marbūṭa : عايزة /ʕajza/, mais elle est remplacée par une 'alif quand elle est suivie par un pronom affixe, ce qui marque l'allongement du /a/ du féminin : هي عايزاه /hijja ʕajza:/ 'elle le veut' (ibid., p. 95).

4.3.2 Orientations Wikipédia maṣri

4.3.2.1 Recommandations officielles

Aucune recommandation officielle relative à la transcription des voyelles longues n'a été trouvée sur Wikipédia maṣri, l'usage semblant ne pas poser de problème.

4.3.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) peut effacer les graphes transcrivant traditionnellement les voyelles longues. Par exemple, il remplace le verbe fuṣḥā جاءوا /dʒa:ʔu:/ par son équivalent 'ammiyya جم /gum/ 'ils sont venus, effaçant ainsi les 'alif et le wāw.

Mais Ghalybot a surtout tendance à rajouter les graphes wāw et yā' dans la translittération de noms étrangers en remplaçant par exemple واشنطن par واشينطون "Washington", أستراليا par أستراليا "Australie", تلفزيون par تليفزيون "télévision", هلسنكي par هيلسينكي "Helsinki". Il semble vouloir transcrire précisément chaque phonème du mot de la langue d'origine par une écriture plus phonographique, en s'éloignant du mode abjad favorisé par l'arabe littéral.

4.3.3 Observations dans le corpus

4.3.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya

Les graphes yā' ي et wāw و sont légèrement plus fréquents sur Wikipédia maṣri (arz) que sur Wikipédia al-'arabiyya (ar), et ce malgré le remplacement recommandé de yā' ي final par 'alif maqṣūra ى. De son côté, la 'alif ا est bien plus fréquente, probablement par effacement de la hamza suscrite ou souscrite, courante dans l'écriture du dialecte.

graphe	Wikipédia maṣri (arz)			Wikipédia al-'arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ي	71.586.498	9,1356%	2	255.872.222	9,1013%	3	3,57
و	46.553.694	5,9410%	6	161.904.072	5,7589%	5	3,48
ا	118.929.723	15,1774%	1	391.184.188	13,9143%	1	3,29
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.16: Fréquence des graphes yā' ي, wāw و et 'alif ا sur arz et ar

4.3.3.2 Exemples de choix graphiques

4.3.3.2.1 Position initiale

En début de mot, les voyelles s'appuient généralement sur le “coup de glotte” /ʔ/, marqué par la 'alif accompagnée, ou non, de la hamza. Cette situation est décrite dans la section 4.4.

4.3.3.2.2 Position médiane

En position médiane, la transcription de /a:/ ou /ɑ:/ long se fait généralement par la 'alif simple ا comme vu avec بابه /ba:ba/ ‘deuxième mois du calendrier copte’ (بابه) et بابا /ba:ba/ ou /b̥a:b̥a/ ‘pape’ (بابا Pape). On trouve même 75 occurrences d'écriture de cette 'alif transcrivant /a:/ là où elle n'apparaît exceptionnellement pas en arabe littéral avec لकिन /la:kin/ ‘mais’ pour la fuṣḥā لكن.

Il a été trouvé plusieurs occurrences d'utilisation de la 'alif pour transcrire l'allongement du /a/ de la tā' marbūṭa suivie d'un pronom affixe dans un participe actif comme pour شافها /ʃajfa:/ ‘elle le voit’ (سيكولاريه) ‘Laïcité’ par Tharwat35), موجهها /muwaggiha:ha/ ‘elle l'oriente’ (سيكولاريه) ‘Statue de la Liberté’ par Ramsis II) et سامعاهم /samʕa:hum/ ‘elle les entend’ (اهل الهوى) par M777). C'est le même phénomène de transcription de la voyelle /a/ allongée que l'on trouve avec معاه /maʕa:/ ‘avec lui’ (848 occurrences dont السعال السودان ‘Soudan du Sud’), معاها /maʕa:ha/ ‘avec elle’ (303 occurrences dont الديكي ‘Coqueluche’), معاهم /maʕa:hum/ ‘avec eux’ (511 occurrences dont كاراكاس ‘Caracas’).

Pour les voyelles /i:/ et /e:/, le recours à yā' ي est d'usage que ce soit pour /i:/ فيل /fi:l/ ‘éléphant’ (فيل), de même prononciation fuṣḥā, ou /e:/ بيت /be:t/ ‘maison’ (بيت), dont la prononciation fuṣḥā est /bayt/. On observe une utilisation occasionnelle des diacritiques

pour différencier d'éventuels homographes comme la kasra pour دِين /di:n/ 'religion' (دين) et la fatha pour دَيْن /de:n/ 'dette'¹⁵.

Pour les voyelles /u:/ et /o:/, le recours à wāw و est d'usage que ce soit pour /u:/ : نور /nu:r/ 'lumière' النور, de même prononciation fuṣḥā, ou /o:/ : نوم /no:m/ 'sommeil' (نوم), dont la prononciation fuṣḥā est /nawm/.

4.3.3.2.3 Position finale

En fin de mot, les voyelles longues de l'arabe littéral s'abrègent en arabe égyptien. Parallèlement, apparaît l'allongement des voyelles finales notamment par affixation du pronom personnel masculin singulier. Pour /i:/, c'est après la deuxième personne du féminin singulier des verbes comme dans متتجوزيه /ma-titgawwizi:/ 'épouse-le' (Farnawany, 1981, p. 168), ou après les prépositions فيه /fi:/ 'il y a' بيه /bi:/ 'avec lui' ليه /li:/ 'il a' (ibid., p. 137). Pour /u:/, c'est après la troisième personne du pluriel des verbes comme dans ع يطمعوه /ʕajtʕammʕu:/ 'il le rendront impatient' (ibid., p. 163). Pour /e:/, ceci concerne les interrogatifs ايه /ʔe:/ 'quoi'¹⁶, ليه /le:/ 'pourquoi' ou la préposition عليه /ʕale:/ 'sur lui' (ibid., p. 126) ou les duels comme /ʕide:/ 'ses deux mains' (ibid., p. 32). Pour /o:/, Farnawany (ibid., p. 126) cite جاتوه /gato:/ 'gâteau' et pour /a:/ كومان سافاه /kumun sava:/ 'comment ça va' (ibid., p. 109). Par ailleurs, Badawi et Hinds (1986, p. 2) écrivent آه pour /a:(h)/¹⁷. L'allongement des voyelles finales est donc transcrit par la cliticisation de ه à ي, ou و, par principe morphologique ou par analogie à celui-ci.

4.4 /ʔ/ ou “coup de glotte” issu de la hamza

4.4.1 État des lieux

4.4.1.1 Situation phonologique

La consonne occlusive glottale sourde [ʔ], à laquelle nous ferons référence par le terme “coup de glotte”, si elle n'est pas issue du recul de l'uvulaire [q] (cf. 4.5), n'est parfois pas prononcée en arabe égyptien.

En position initiale, le phonème [ʔ] est généralement conservé dans les pluriels du type /ʔasʕha:b/ 'amis', mais la syllabe dont il est l'attaque peut disparaître dans /ʔakal/ qui devient /kal/ 'manger', et /ʔaxad/ qui devient /xad/ 'prendre'. Dans une suite de mots proches ou prononcés rapidement, [ʔ] initial s'élide régulièrement : /ʕiribt ʔana/ > /ʕiribt ana/ 'c'est moi qui ai bu'. Abdel-Massih, Abdel Malek et Badawi (2009, p. 103-104) différencient : /ʕawz adxul/ 'je (masculin) veux entrer' de /ʕawza ʔadxul/ 'je (féminin) veux entrer' car «l'identité d'un mot ou d'une expression serait obscurcie» si [ʔ] était élide. De même, l'élision n'a pas lieu si cela implique la rencontre de deux voyelles longues : /li: ʔa:xir/ 'cela a une fin'.

¹⁵C'est le même usage qu'en fuṣḥā avec utilisation de la kasra pour دِين /di:n/ 'religion' et de la fatha pour دَيْن /dajn/ 'dette'.

¹⁶Apparenté à la fuṣḥā آيه /ʔajju-hu/ 'lequel' (Farnawany, 1981, p. 141).

¹⁷Badawi et Hinds (1986) transcrivent /ʔe:h/, /le:h/, /gato:h/ là où Farnawany (1981) omet le /h/, reflétant ainsi la prononciation courante.

En position médiane, à côté de /raʔs/ (considéré comme relevant du registre élevé de la langue), on entend /ra:s/ pour ‘tête’ (Badawi et Hinds, 1986, p. 319), dans un phénomène d’allongement compensatoire de la voyelle courte précédent [ʔ]. À la place du fuṣḥā /biʔr/, on entend /bi:r/ pour ‘puits’ (ibid., p. 115). Par ailleurs, les participes actifs des verbes trilitères de forme 1, dont la radicale médiane est [ʔ] en fuṣḥā, voient celle-ci transformée en semi-voyelle [j] : fuṣḥā : /na:ʔim/ = ‘āmmiyya : /na:jim/ ‘dormant’ (Gadalla, 2000, p. 33). Cette transformation est d’ailleurs observable en arabe littéral parmi les verbes dont la radicale finale est [ʔ], avec à côté de la hamza de قراءة /qira:ʔa/ ‘lecture’ on trouve le yā’ de بداية /bida:ja/ ‘début’ (Badawi, Carter et Gully, 2016, p. 13).

En position finale, la disparition de la hamza en arabe égyptien est quasiment systématique, sans allongement compensatoire : fuṣḥā : /sama:ʔ/ = ‘āmmiyya : /sama/ ‘ciel’, fuṣḥā : /malaʔ/ = ‘āmmiyya : /mala/ ‘remplir’ (Gadalla, 2000, p. 30). Pour les verbes, après ajout d’un suffixe commençant par une consonne, il y a allongement compensatoire : fuṣḥā : /malaʔtu/ = ‘āmmiyya : /male:t/ ‘j’ai rempli’. L’aboutissement à la voyelle /e:/ implique le passage par la suite /aj/ (ibid., p. 72).

4.4.1.2 Réalisation graphique

En position initiale, Badawi et Hinds (1986, p. 9) transcrivent le “coup de glotte” par la hamza, selon les règles fuṣḥā, c’est-à-dire en plaçant celle-ci sous une ’alif comme dans إحنا /iḥna/ ‘nous’ pour /ʔi/, ou sur une ’alif comme dans أخ /aḫ/ ‘frère’ pour /ʔa/ ou dans أخت /aḫt/ ‘sœur’ pour /ʔu/. Pour les mots commençant par deux consonnes consécutives, et donc précédées à l’écrit en fuṣḥā par la ’alif surmontée de la hamzat al-waṣl, ils respectent la convention fuṣḥā en écrivant une ’alif simple, tout en indiquant la prononciation en début de phrase, c’est-à-dire avec le “coup de glotte”, comme dans ابن /ibn/ ‘fils’ Badawi et Hinds (ibid., p. 5), اثنين /itne:n/ ‘deux’ Badawi et Hinds (ibid., p. 139), است /ist/ ‘fesses’ Badawi et Hinds (ibid., p. 18), اسم /ism/ ‘nom’ Badawi et Hinds (ibid., p. 23), الله /alla:h/ ‘Dieu’, اللى /illi/ ‘qui, que (pronom relatif)’ Badawi et Hinds (ibid., p. 33), اتولد /itwalad/ ‘il est né’ Badawi et Hinds (ibid., p. 955), etc. En milieu de phrase, le signe /ʔ/ tombe dans la transcription, selon les règles de prononciation fuṣḥā et ‘āmmiyya, comme dans /w-alla:hi/ ‘par Dieu’ (ibid., p. 34) ou /ʔis-sitt illi sakna ʃan nasʕja/ ‘la femme qui vit au coin de la rue’. On trouve cependant aussi /illi/ en milieu de phrase comme dans /ʔil-ḥaʔʔa ʃalajja ʔilli tawiʕt-ak/ ‘c’est ma faute, [c’est moi] qui t’ai obéi’ (ibid., p. 33). On trouve également la ’alif madda ʾ pour transcrire /ʔa:/ en début de mot comme pour آدم /a:dam/ ‘Adam’ (ibid., p. 1), mais aussi pour transcrire /ʔa/ comme dans آلو /alo:h/ ‘allô’ (ibid., p. 34), à côté de ألاى /ala:j/ ‘régiment’ (ibid., p. 31). De même, pour /ʔu/ initial on trouve la ’alif mahmūza suivie de wāw comme dans أوبرا /ubra/ ‘opéra’ (ibid., p. 3) et أوتوستوب /utustubb/ ‘autostop’ à côté de la ’alif mahmūza seule comme dans أتمبيل /utumbil/ ‘automobile’ (ibid., p. 6). Pour /ʔi/ initial on trouve on trouve la ’alif mahmūza suivie de yā’ comme dans إيطاليا /itʕalja/ ‘Italie’ à côté de la ’alif mahmūza seule comme dans إطار /itʕa:t/ ‘pneu’ (ibid., p. 26). La transcription du “coup de glotte” en début de mot se fait donc essentiellement selon les règles fuṣḥā, avec une tendance à transcrire la voyelle courte suivante par un graphe traditionnellement associé aux voyelles longues.

Farnawany (1981, p. 71) explique que, quand le “coup de glotte” initial est éliminé par déplacement en position médiane, une ‘alif¹⁸ est écrite étymologiquement par certains écrivains. Il écrit et transcrit *جوز اختها* /go:z uxtaha/ ‘le mari de sa soeur’ *بكره اخدها* /bukra xudha/ ‘demain je la prends’. D’autres écrivains adoptent, mais très rarement, une orthographe phonémique où ils omettent cette ‘alif : *ونا كان* /w-ana kama:n/ ‘et moi aussi’ (au lieu de *وانا كان*), *لجلي* /lagli/ ‘pour moi’ (au lieu de *لاجلي*), etc.

En position médiane, quand le “coup de glotte” fuṣḥā suivi de la voyelle /i/ devient /j/ en ‘āmmiyya, Rosenbaum (2004, p. 293) indique que la hamza fuṣḥā est alors remplacée par yā’ comme quand *الجائزة* devient *الجايذة* /il-gajza/ ‘le prix, la récompense’, quand *زبائن* devient *زباين* /zaba:jin/ ‘clients’, etc. Quand le “coup de glotte” fuṣḥā directement suivi d’une consonne, sans voyelle intermédiaire, s’élide avec allongement de la voyelle précédente, il indique que la hamza fuṣḥā est alors remplacée par la voyelle longue correspondante comme quand *بئر* devient *بير* /bi:r/ ‘puits’, quand *رأسي* devient *راسي* /ra:si/ ‘ma tête’, quand *الشوم* devient *الشوم* /iʃ-fu:m/ ‘l’indigne (personne)’.

En position finale, quand la séquence fuṣḥā /a:ʔ/ devient /a/ en ‘āmmiyya, Rosenbaum (ibid., p. 293) indique que la hamza est souvent effacée comme quand *خضراء* devient *خضرا* /xadʕra/ ‘verte’, quand *خبراء* devient *خبرا* /xubara/ ‘experts’, quand *سواء* devient *سوا* /sawa/ ‘ensemble’, quand *عشاء* devient *عشا* /ʕaʃa/ ‘dîner’, etc. Farnawany (1981, p. 73) précise que, pour cette graphie phonémique, la ‘alif ا peut être remplacée par la ‘alif maqṣūra ي et il relève dans son corpus *سوى* /sawa/ ‘ensemble’ et *العشى* /il-ʕaʃa/ ‘le dîner’. Il remarque que, particulièrement dans cette position finale, la graphie étymologique fuṣḥā est souvent conservée, comme dans *الصحراء* /isʕ-sʕaħra/ ‘le désert’, *سيناء* /si:na/ ‘le Sinaï’ ou *عشاء* /ʕaʃa/ ‘dîner’.

4.4.2 Orientations Wikipédia maṣri

4.4.2.1 Recommandations officielles

Dans les recommandations qui figurent sur la page officielle Ṭarīqa al-kitāba طريقة الكتابة ‘méthode d’écriture’ (Wikipédia, 2025c) le premier élément concerne la hamzat al-qaṭ’, essentiellement son utilisation sur ou sous ‘alif, la règle étant d’écrire ‘alif sans hamza (ni dessus, ni dessous) : /ʔaw/ او ‘ou’ et /ʔilla/ الا ‘sauf’, par exemple, étant de fait majoritairement écrits sans hamza dans la page « méthode d’écriture ». Les exceptions présentées pour lesquelles il est recommandé de conserver la hamza sont pour

- les mots qui commencent avec ‘alif avec hamza et lām comme ألف /ʔælf/ ‘mille’ ou إلف /ʔelʃ/ ‘Elvis’, ou avec ‘alif avec madda et lām comme آل /ʔæl/ ‘parenté’, pour différencier d’un mot déterminé par l’article lām al-taʕrīf : الحرف /ʔi)l-ħarf/ ‘la lettre’ par exemple, celui-ci commençant par une ‘alif simple (théoriquement avec hamzat al-waṣl).
- les mots dont la ‘alif avec hamzat al-qaṭ’ est au milieu : تأمين /taʔmi:n/ ‘assurance’

¹⁸Il écrit : “Trotz dieser Elision wird - ء - von manchen Schriftsteller etymologisch geschrieben, von anderen dagegen weggelassen.” (Malgré cette élimination, - ء - est écrit étymologiquement par certains auteurs, tandis que d’autres l’omettent) mais ses exemples (manuscrits) n’utilisent que des ‘alif sans hamza. Il ne semble pas différencier celle-ci de la ‘alif avec hamza.

- les autres formes de hamza : مؤلف /muʔallif/ 'auteur', شيء /ʃe:ʔ/ 'chose', بدائي /bi-da:ʔi/ 'primitif'.

Samsam22 exprime le 31/12/2008 (archive 2) son désaccord avec la règle de l'effacement de la hamza avec 'alif, arguant du fait que sa position au dessous ou au dessus permet de différencier entre la kasra prononcée /i/ et la fatha prononcée /a/ (sans mentionner la damma prononcée /u/). La formulation de la règle d'origine n'a cependant pas été changée.

4.4.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) efface la hamza placée sur ou sous la 'alif initiale des mots suivants : أبريل "avril", أغسطس "août", أكتوبر "octobre", أفريقيا "Afrique", أمريكا "Amérique", أوروبا "Europe", إثيوبيا "Ethiopie", أرمينيا "Arménie", ألمانيا "Allemagne", إنجلترا "Angleterre", إيطاليا "Italie", إسبانيا "Espagne", أستراليا "Australie", الأردن "Jordanie", الأمم المتحدة "Les Nations Unies", أكثر "plus", أكبر "plus grand". Il change أحسن en أفضل "mieux" et أعطى en ادا "donner". Il efface la madda آ sur la 'alif initiale de آيسلندا "Islande".

Il transforme la hamza médiane يـ en يā dans دائم "permanent", دائرة "cercle", سائح "touriste", gagnant "gagnant", جائزة "récompense", جوائز "récompenses", خسائر "pertes", رسائل "messages", طوائف "sectes", ملائكة "anges", هزائم "défaites", وسائل "moyens", وظائف "emplois", les mariés "العرائس", بئر نفط "puits de pétrole" et transforme عائلة en عيلة "famille".

Il efface la hamza finale de زملاء "collègues", علماء "savants", المحيط الهادئ "Océan Pacifique" et change ابتدا en بدء "commencer" et قراءة en قرأة "lecture".

Il conserve la hamza dans الأشهر "le plus célèbre", إن "que", أتأسس "être fondé", متأثر "être influencé", رئيس "président", مسؤول "responsable", مؤرخ "historien", جزء "partie".

Les modifications opérées par Ghalybot respectent donc les recommandations de la page officielle.

4.4.3 Observations dans le corpus

4.4.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya

Comme observé dans le tableau 4.17, les différentes graphies de la hamza ء, ؤ, و, ainsi que celles utilisant la 'alif comme support : إ, أ, ي, compris la madda آ, sont relativement moins utilisées sur Wikipédia maṣri par rapport à Wikipédia al-'arabiyya. Cela correspond à la quasi absence de prononciation du "coup de glotte" /ʔ/ en milieu et fin de mot en arabe égyptien, et donc à l'effacement de la graphie étymologique au profit d'une graphie plus phonologique.

Concernant la 'alif, la règle enjoignant à l'écrire sans hamza (ni dessus, ni dessous) semble expliquer la relative rareté de إ et أ sur arz par rapport à ar et par conséquent la fréquence légèrement plus élevée à l'inverse pour la 'alif "nue" ا sur Wikipédia maṣri en comparaison avec Wikipédia al-'arabiyya. Si on considère le total des occurrences de ces graphes sur chacune des deux éditions Wikipédia, on obtient un ratio de 3,83 occurrences sur ar pour chaque occurrence sur arz, proche du ratio de 3,59 entre les totaux des occurrences de l'ensemble des caractères étudiés dans le corpus.

graphie	Wikipédia maṣri (arz)			Wikipédia al-ʿarabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ا	348.229	0,0444%	36	19.730.202	0,7018%	25	56,66
أ	2.275.450	0,2904%	28	54.292.275	1,9312%	16	23,86
ؤ	152.032	0,0194%	39	2.868.430	0,1020%	36	18,87
آ	221.304	0,0282%	37	2.488.371	0,0885%	38	11,24
ئ	1.573.673	0,2008%	29	94.74.864	0,3370%	32	6,02
ء	1.268.363	0,1619%	30	6.591.327	0,2345%	33	5,20
إ	118.929.723	15,1774%	1	391.184.188	13,9143%	1	3,29
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.17: Fréquence de la hamza et madda sur arz et ar

4.4.3.2 Exemples de choix graphiques

Dans Wikipédia maṣri, le “coup de glotte” qui subsiste de l’arabe littéral suit généralement les règles orthographiques fuṣḥā, sauf à l’initiale où la recommandation est de ne pas la marquer.

Pour beaucoup de mots, le “coup de glotte” /ʔ/ initial reste indiqué par le diacritique hamza ʾ pour /ʔa/ ou /ʔa/ (tableau 4.18), ou madda ̣ pour /ʔa:/ ou /ʔa:/ (tableau 4.19). Dans les exemples testés, le diacritique fatha ˆ n’a jamais été vu sur la hamza.

traduction	API	ا			أ			total
			#	%		#	%	
plus long	/ʔatˤwa:l/	اطول	1.574	81	أطول	354	18	1.928
blanc	/ʔabjadˤ/	ابيض	4.423	75	أبيض	1.414	24	5.837
bleu	/ʔazraʔ/	ازرق	1.137	64	أزرق	622	35	1.759
enfants	/ʔatˤfa:l/	اطفال	4.522	74	أطفال	1.509	25	6.031
ailes	/ʔagniħa/	اجنحه	2.199	99	أجنحه	15	01	2.214
		اجنحة	105	48	أجنحة	112	51	217
total			13.960	77		4.026	22	17.986

Tab. 4.18: /ʔa/ ou /ʔa/ initial transcrit par ا ou أ

Pour les mots corrigés massivement par Ghalybot, cette recommandation est respectée, comme on observe dans le tableau 4.20. Les formes sans hamza sur la ’alif sont prépondérantes, mais on remarque tout de même une forte résistance de أكثر (avec le graphe ث fuṣḥā), ce qui relativise l’influence de ce “bot”.

Mais ce programme ne fait que renforcer une tendance qui semble naturelle chez les rédacteurs de Wikipédia maṣri, bien que moins marquée, comme illustré dans le tableau 4.21.

Sans intervention de Ghalybot, l’effacement de la hamza est moins systématique. Par exemple, celle-ci reste hégémonique sur la ’alif de /ʔaw/ ‘ou’ avec 237.208 occurrences

traduction	API	ا			آ			total
			#	%		#	%	
Adam	/ʔa:dam/	ادم	422	40	آدم	627	59	1.049
Oui / Ah !	/ʔɑ:(h)/	اه	242	38	آه	393	61	635
total			664	39		1.020	60	1.684

Tab. 4.19: /ʔa:/ ou /ʔɑ:/ initial transcrit par ا ou آ

traduction	API	sans hamza			avec hamza			total
			#	%		#	%	
Arménie	/ʕarmi:nja/	ارمينيا	4.676	87	أرمينيا	661	12	5.337
prix	/gawa:jiz/	جوايز	6.272	77	جوائز	1.777	22	8.049
famille	/ʕe:la/	عيلة	3.237	83	عائلة	658	16	3.895
		عيله	3.315	73	عائله	1.176	26	4.491
savants	/ʕulama/	علماء	16.983	95	علباء	860	04	17.843
plus	/ʔaktar/	اكتر	533	35	أكثر	956	64	1.489
		اكثر	8.937	89	أكتر	1.024	10	9.961
plus grand	/ʔakbar/	اكبر	13.998	96	أكبر	494	03	14.492
total			57.951	88		7.606	11	65557

Tab. 4.20: Emploi ou omission de hamza initiale (Ghalybot)

de او contre seulement 22.633 أو. La concision de ce mot explique peut-être la nécessité de bien l'identifier dans son contexte.

L'écriture de /ʔu/ initial, comme illustré dans le tableau 4.22 se fait très majoritairement sans hamza sur la 'alif initiale, pour les mots remplacés par Ghalybot comme /ʔusturalja/ et /ʔurubba/, comme pour les autres mots testés.

On observe une hésitation pour la transcription de /ʔusturalja/ 'Australie' entre l'utilisation en début de mot de la 'alif (avec ou sans hamza) seule ou suivie du graphe wāw و, bien que la voyelle /u/ ne soit pas prononcée longue. C'est pourtant cette dernière solution qui est largement préférée.

Pour ce qui est de /ʔi/ ou /ʔe:/ initial, le tableau 4.23 illustre les tendances d'écriture de la hamza sous la 'alif. Avec les échantillons choisis, celle-ci est rarement marquée. On trouve même la hamza sur la 'alif, en contradiction avec la règle exigeant de la placer en dessous si sa voyelle est /i/. Ici encore, les remplacements effectués par Ghalybot pour /ʔitʕalja/ 'Italie' et /ʔingiltira/ 'Angleterre' semblent imposer l'hégémonie de la 'alif sans hamza. On remarque aussi l'emploi de la suite 'alif puis graphe yā ي pour transcrire /ʔi/ court initial, comme pour إيطاليا 'Italie', phénomène que l'on retrouve avec ايرلندا 'Irlande' (prononcée /ʔirlanda/ ou /ʔajirlanda/) en contraste avec ايران /ʔi:ra:n/ 'Iran' dont le graphe yā correspond bien à une voyelle longue (Wehr et Cowan, 1979, p. 47).

En milieu de mot, la hamza se transforme régulièrement en 'āmmiyya en la voyelle longue qu'elle porte en fuṣḥā. C'est ce que l'écriture de la majorité des occurrences de /bi:r/ 'puits' et /ra:s/ 'tête' reflète sur Wikipédia maṣri, comme l'illustre le tableau 4.24. Cependant, l'écriture avec hamza reste assez courante, ce qui peut aussi correspondre à

traduction	API	sans hamza			avec hamza			total
			#	%		#	%	
plus petit	/ʕasˤɣar/	اصغر	892	46	أصغر	1.043	53	1.935
moins	/ʕaʕall/	اقل	337	30	أقل	753	69	1.090
Albanie	/ʕalba:nja/	البانيا	6.194	96	ألبانيا	243	03	6.437
crimes	/gara:jim/	جرائم	984	78	جرائم	267	21	1.251
cercles	/dawa:jir/	دوائر	90	27	دوائر	238	72	328
poètes	/ʃuʕara/	شعرا	520	55	شعراء	416	44	936
total			9.017	75		2.960	24	11.977

Tab. 4.21: Emploi ou omission de hamza initiale (hors Ghalybot)

traduction	API	ا			أ			total
			#	%		#	%	
Australie	/ʔusturalja/	استراليا	6.543	97	أستراليا	166	02	6.709
		اوستراليا	60.202	99	أوستراليا	5	00	60.207
Europe	/ʔurubba/	اوروبا	25.831	95	أوروبا	1.084	04	26.915
professeur	/ʔusta:z/	استاذ	14.115	92	أستاذ	1.184	07	15.299
origines	/ʔusˤu:l/	اصول	525	59	أصول	352	40	877
total			107.216	97		2.791	02	110.007

Tab. 4.22: /ʔu/ initial transcrit par ا ou أ

une conservation de la prononciation fuṣḥā.

En milieu de mot, la séquence /a:ʔi/ que l'on trouve dans les participes actifs de type /na:ʔim/ 'dormant' ou les pluriels /xasa:ʔir/ 'pertes' en fuṣḥā devient généralement /a:ji/ en ʿammiyya. Sur Wikipédia maṣri, le tableau 4.25 montre une légère préférence pour l'adoption de l'écriture phonémique avec yā ي, sans effets notoires des remplacements effectués par Ghalybot pour /dajra/ et /wazˤa:jif/.

Pour les cas, en milieu de mot, où différentes règles d'écriture de la hamza sont en concurrence, les rédacteurs de Wikipédia maṣri suivent massivement la règle de base et respectent peu les recommandations de l'Académie de la langue arabe du Caire qui préconise d'éviter la suite de deux lettres semblables comme و و (voir page 41), ainsi que l'attestent les statistiques du tableau 4.26. C'est cette suite de deux graphes similaires qui est plébiscitée par les rédacteurs, tout comme sur Wikipédia al-ʿarabiyya où les taux d'emploi sont de 98% pour مسؤول et رؤوس et de 89% pour شؤون.

En finale de mot, le "coup de glotte" a aussi tendance à disparaître en arabe égyptien, et cela se traduit à l'écrit par la disparition de la hamza, comme illustré par les résultats présentés dans le tableau 4.27. La tendance est effectivement à ne pas écrire la hamza quand elle n'est pas prononcée.

Comme mentionné par Panović (2010, p. 109-110), il existe de rares exemples d'hypercorrection avec l'orthographe قوده employée par le rédacteur Samsam22¹⁹ en lieu et

¹⁹ dans الدُر 'Chajar ad-Durr', شامبوليون 'Jean-François Champollion' et محاكم التفتيش 'Inquisition'

traduction	API	ا			أ			إ			total
			#	%		#	%		#	%	
toi	/ʔint-/	انت	1.193	74	أنت	354	21	إنت	65	04	1.612
toi (sm)	/ʔinta/	انت	5	45	أنت	6	54	إنت	0	0	11
toi (sf)	/ʔinti/	انت	3	21	أنت	10	71	إنت	1	07	14
		انتِي	44	61	أنتِي	16	22	إنتِي	11	15	71
		انتى	145	88	أنتى	7	04	إنتى	12	07	164
quoi	/ʔe:/	ايه	4.029	90	أيه	50	01	إيه	380	08	4.459
Italie	/ʔitˤalja/	إيطاليا	87.011	99	أيطاليا	5	00	إيطاليا	218	00	87.234
Angleterre	/ʔingiltira/	انجلترا	19.385	99	أجلترا	6	00	إنجلترا	161	00	19.552
total			111.815	98		454	00		848	00	113.117

Tab. 4.23: /ʔi/ initial transcrit par ا, أ ou إ

traduction	API	ا / ي			أ / ئ			total
			#	%		#	%	
puits	/bi:r/	بير	745	60	بئر	477	39	1.222
tête	/ra:s/	راس	1.067	64	رأس	588	35	1.655
total			1.812	62		1.065	37	2.877

Tab. 4.24: Hamza médiane conservée ou non

place des graphies plus habituelles أودة / أوضة pour /ʔo:dˤa/ ‘pièce, chambre’ apparenté au turc “oda” de même sens (Badawi et Hinds, 1986, p. 43). Ceci laisse à penser que qāf est devenu le graphe privilégié par certains rédacteurs pour la transcription du “coup de glotte” à la place de la hamza, au moins en position initiale pour les mots d’origine étrangère.

4.5 /ʔ/ ou “coup de glotte” issu de /q/

4.5.1 État des lieux

4.5.1.1 Situation phonologique

En fuṣḥā, le graphe ق transcrit la consonne occlusive uvulaire sourde [q]. En dialecte du Caire (et au Levant), ce phonème est généralement remplacé par son équivalent glottal [ʔ], parfois appelé “coup de glotte”, phonème qui existe aussi indépendamment en ‘ām-miyya comme en fuṣḥā, où il est transcrit par la hamza ء (cf. 2.2). Ceci peut donner lieu à des prononciations ambiguës en ‘ām-miyya, comme illustré dans le tableau 4.28.

Le phonème [q] est conservé en arabe égyptien dans un petit nombre de mots, souvent associés au niveau diglossique considéré comme élevé²⁰ : /ʔilqa:hira/ ‘Le Caire’, /qa:hira/

²⁰Selon Farnawany (1981, p. 17), dans la langue des gens sans instruction qui tentent de prononcer /q/, à la

traduction	API	ي			ئ			total
			#	%		#	%	
dormant	/na:jim/	نايم	43	84	نائم	8	15	51
cercle	/dajra/	دايرة	359	72	دايرة	138	27	497
pertes	/xasa:jir/	خسائر	218	41	خسائر	302	58	520
fonctions	/wazʕa:jif/	وظائف	961	55	وظائف	762	44	1.723
total			1.581	56		1.210	43	2.791

Tab. 4.25: /ji/ médian transcrit par ي ou ئ

	responsable	affaires	têtes	
	/masʔu:l/	/ʃuʔu:n/	/ruʔu:s/	
ؤو	مسؤول	شؤون	رؤوس	Total
	9.517	3.734	182	13.433
	93%	83%	98%	90%
ئو	مسؤول	شؤون	رئوس	Total
	589	709	1	1.299
	6%	16%	1%	9%
ؤ	مسؤول	شؤون	رؤوس	Total
	106	29	0	135
	1%	1%	0%	1%
ءو	مسؤول	شؤون	رءوس	Total
	0	0	2	2
	0%	0%	1%	0%
Total	10.212	4.472	185	14.869
	100%	100%	100%	100%

Tab. 4.26: /ʔu:/ médian transcrit par hamza sur différents supports

/qahri/ ‘irréfutable’ (Badawi et Hinds, 1986, p. 720), /ʔilqurʔa:n/ voire /ʔilqura:n/ ‘le Coran’, /qarya/ (pluriel /qura/) ‘village’, /qazir/ /ʔaqza:r/ ‘sale, plus sale’ (ibid., p. 698), /qardʕ/ ‘prêt’ (ibid., p. 694), /ʔiqtarah/ ‘suggérer’, /qara:r/ ‘décision’, /qarra/ ‘continent’ (ibid., p. 692), /qisʕa:sʕ/ peine (pendaison pour meurtre) (ibid., p. 704), /ʔaqalliyya/ ‘minorité’ (ibid., p. 716), /musi:qa/ ‘musique’ (ibid., p. 824), /asqa:l/ ‘haltères’ (ibid., p. 131), /raqam/ ‘nombre’ (ibid., p. 348), /saqa:fa/ ‘culture’ (ibid., p. 418), etc.

On trouve aussi des alternances voire des paires minimales citées par Al-Tonsi, Al-Sawi et Massoud (2010, p. 4) et illustrées dans le tableau 4.29.

Il est à noter que dans le régiolecte ʕaʕdi (de Haute Égypte) et dans celui de la partie est du Delta la prononciation de qāf ق est généralement le /g/ (et celle de ġīm ج est le /dʒ/ et non le /g/ cairote) (Manfred Woidich, 2008, p. 7) (Khalafallah, 1969, p. 41). Ainsi جوقة ”troupe musicale” est prononcé /dʒawqa/ en fuṣḥā (Wehr et Cowan, 1979, p. 176), /go:ʔa/ en dialecte du Caire (Badawi et Hinds, 1986, p. 183) par postériorisation de l’uvulaire vers

place de la valeur sonore [q], apparaît souvent une variante [kʕ], qui se prononce plus en arrière que /k/ normal et avec emphase. Il préfère donc transcrire /kʕurʔa:n/ ‘Coran’ - /ilkʕa:hirʕa/ ‘Le Caire’ - /sakʕa:fa/ ‘culture’ (Farnawany, 1981, p. 110).

traduction	API	Ø			ء			total
			#	%		#	%	
savants	/ʕulama/	علما	928	69	علباء	415	30	1.343
poètes	/ʕuʕara/	شعرا	1.845	72	شعراء	716	27	2.561
ciel	/sama/	سما	90	41	سماء	126	58	216
derrière	/wara/	ورا	1.299	73	وراء	457	26	1.756
calme	/ha:di/	هادى	1.819	53	هادئ	436	13	3.437
		هادي	1.182	34				
total			7.163	76		2.150	23	9.313

Tab. 4.27: Hamza finale conservée ou non

fuṣḥā		‘āmmiyya	traduction
قلم	/qalam/	/?alam/	crayon (Badawi et Hinds, 1986, p. 716)
ألم	/ʔalam/		douleur (Badawi et Hinds, 1986, p. 33)

Tab. 4.28: Passage de /q/ en fuṣḥā à /ʔ/ en ‘āmmiyya

la glottale et /dʒo:ga/ en régiolecte ʕaʕdi par antériorisation de l’uvulaire vers la vélaire et voisement.

4.5.1.2 Réalisation graphique

Farnawany (1981) et Rosenbaum (2004) s’accordent pour constater que le phonème /ʔ/ ou “coup de glotte” issu de /q/ est presque toujours transcrit par le graphe ق, comme en fuṣḥā, et rarement phonémiquement par la hamza ʔ. Ils citent cependant quelques cas de ces exceptions : أد /ʔadd/ ‘combien’, أوى /ʔawi/ ‘fort’, أدام /ʔudda:m/ ‘devant’, أمورة /ʔammu:ra/ ‘belle’, etc. (Farnawany, 1981, p. 70-71), ou encore اطر /ʔatʕr/ ‘train’, اوي /ʔa-wi/ ‘très’, ازرق /ʔazraʔ/ ‘bleu’, etc. (Rosenbaum, 2004, p. 286). Farnawany (1981, p. 70) évoque l’écriture dialectale expérimentée par Maḥmūd Taymūr (1894-1973) dans plusieurs de ses pièces de théâtre où, à côté de la version fuṣḥā, il a publié une version ‘āmmiyya dans laquelle pratiquement tous les “coup de glotte” étaient remplacés par une forme de la hamza, et on trouve ainsi تداي /tida:jiʔ/ ‘qui irrite’, وأف /waʔf/ ‘arrêt’, دئيته /diʔi:ʔa/ ‘minute’, إلحوتنا /ilḥaʔu:na/ ‘sauvez-nous’, etc. (Taymūr, 1949). Cette expérimentation réalisée dans un esprit de nationalisme égyptien a été abandonnée par Taymūr à l’heure du panarabisme triomphant et qualifiée de “réalisme pur, limité et superficiel”²¹ (Diem, 2006, p. 114). Avallone (2015, p. 82) observe dans la troisième saynète de son échantillon de textes extraits du magazine *al-ʿUstād* de ʿAbdallāh an-Nādim (1842-1896) un grand nombre de mots contenant un coup de glotte qui, dans l’écriture arabe fuṣḥā, correspond au graphème ق et qui sont écrits dans ce texte avec une hamza. Elle cite ارش /ʔirʃ/ ‘piastre’, لؤمه /luʔma/ ‘bouchée (de pain)’, وائف /wa:ʔif/ ‘debout’, illustrant ainsi la diversité des supports de la hamza utilisés pour transcrire la prononciation ‘āmmiyya, avec probablement une intention parodique. La chercheuse n’a pas observé de phénomène

²¹ “reinen, beschränkten und oberflächlichen Realismus”

fuṣḥā		‘āmmiyya	traduction
مُسْتَقِلّ	/mustaqill/	/mistiʔill/	qui trouve que c’est peu
		/mustaqill/	indépendant
قَرْن ج قُرُون	/qarn/ pl. /quru:n/	/ʔarn/ pl. /ʔuru:n/	cosse, corne
		/qarn/ pl /quru:n/	siècle
قسم ج أقسام	/qism/ pl. /ʔaqsɑ:m/	/ʔism/ pl. /ʔaʔsa:m/	poste de police
		/qism/ pl. /ʔaqsɑ:m/	subdivision, section
قانون	/qa:nu:n/	/ʔanu:n/	qānūn (instrument)
		/qanu:n/	loi

Tab. 4.29: Paires minimales distinguées par /q/ ou /ʔ/

similaire dans ses autres échantillons du xix^e siècle.

4.5.2 Orientations Wikipédia maṣri

4.5.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle Ṭarīqa al-kitābah طريقة الكتابة ‘méthode d’écriture’ (Wikipédia, 2025c) comportent l’élément suivant :

ق
لو فيه كلمه من أصل عربي فيها حرف "ق", يتساب زى ما هو.
”Si un mot d’origine arabe contient la lettre ق, elle reste telle quelle.” (Es-perfulmo, le 02/04/2009).

Ceci implique que l’orthographe étymologique doit être conservée même si la réalisation phonologique de ce graphe est /ʔ/ en parler du Caire ou /q/ en parler ša’īdi, et non /q/, la prononciation traditionnelle en arabe littéral.

Le remplacement par une hamza n’est donc pas recommandé, bien qu’il ait été envisagé par un contributeur anonyme dans la page de discussion (archive 3) :

القاف تبدل بالهمزة في كل الكلمات المصري طالمه هاتتنطق همزة. امثلة: بتالك هنا أد إيه
(مش: بقالك هنا قد إيه).
”Le qāf est remplacé par la hamza dans tous les mots égyptiens tant qu’il se prononce [comme une] hamza. Exemples : ça fait combien de temps que tu es ici ?²²”

On trouve aussi la graphie قوم ووم pour /ʔu:m/ ‘lève-toi’ (archive 3) et حقوق حأوأ pour /huʔu:ʔ/ ‘droits’ (archive 1). Dudi (archive 1) propose le 19/10/2008 de différencier entre قوي ‘fort’ et اوي ‘très’ par exemple dans الرجل ده قوي أوي /ir-ra:gil da ʔawi ʔawi/ ‘cet homme est très fort’, bien que, selon Badawi et Hinds (1986, p. 725), dans les deux mots le qāf classique se prononce /ʔ/ (sauf quand /qawi/ signifie ‘Tout-Puissant’).

Un contributeur anonyme (archive 3) propose de faire la différence entre ق prononcé /q/ القانون فوء الجميع /il-qanun foʔ il-gami:ʕ/ ‘la loi est au dessus de tous’ et celui prononcé

²²/baʔallak hina ʔadd ʔe:/

/ʔ/ qui serait remplacé par une hamza : *الـثانـون آلة موسيقية* : /il-ʔanun ʔa:la musiɟijja/ ‘le qānūn est un instrument de musique’.

Le fait est que le recours à la hamza n’a pas été retenu dans la politique officielle. Sans que cela soit explicité, cette solution permet de laisser la possibilité aux locuteurs du régiolecte ʕaʔdi de lire ق avec leur prononciation [q] traditionnelle.

4.5.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) suit globalement les recommandations et remplacent certaines expressions par d’autres qui conservent le graphe ق bien qu’il corresponde à la prononciation /ʔ/ en parler du Caire : قبل ‘avant que’, دلوڤتي ‘maintenant’, حقيقي ‘vrai’, وقت ‘temps’, شرق ‘est’, فريق ‘équipe’, قرايب ‘proches’, مناقشات ‘discussions’, قراءة ‘lecture’, قهاوى ‘cafés’, سواق ‘chauffeur’, etc. Par contre, il remplace زجاجة par ازازة au lieu de قزازة pour ‘bouteille’ (Badawi et Hinds, 1986, p. 698), ainsi que جدا ‘très’ et للغاية ‘extrêmement’ par اوى au lieu de قوى pour ‘très’, marquant ainsi la prononciation cairote du qāf.

4.5.3 Observations dans le corpus

4.5.3.1 Fréquence d’utilisation des graphes par rapport à Wikipédia al-ʕarabiyya

Comme le montre le tableau 4.30, le graphe ق est presque aussi fréquent sur Wikipédia maʕri (arz) que sur Wikipédia al-ʕarabiyya (ar). Cela semble confirmer le manque de popularité de la solution phonétique de la hamza pour marquer le “coup de glotte” /ʔ/, qui remplace en arabe du Caire le /q/ de l’arabe fuṣḥā, et la forte tendance à conserver la graphie étymologique.

graphe	Wikipédia maʕri (arz)			Wikipédia al-ʕarabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ق	12.080.564	1,5417%	18	52.283.490	1,8597%	17	4,33
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.30: Fréquence du graphe ق sur arz et ar

4.5.3.2 Exemples de choix graphiques

Le choix des rédacteurs des articles de Wikipédia maʕri est très majoritairement de marquer le “coup de glotte” par ق, préférant ainsi une orthographe étymologique correspondant à l’usage phonémique en fuṣḥā, conformément aux recommandations orthographiques de la page officielle sur la méthode d’écriture (Wikipédia, 2025c). Cela permet accessoirement de laisser au lecteur le choix de la prononciation effective, selon le régiolecte préféré. On trouve ainsi les articles dont le titre est قلب ‘coeur’ qui peut être lu /ʔalb/ en parler du Caire ou /galb/ en parler ʕaʔdi. De même, le titre de l’article دقيقة ‘minute’ peut être lu /diʔi:ʔa/ ou /dagi:ga/.

Cependant, on trouve aussi l’article intitulé اطر /ʔaʔr/ ‘train’, alors que le mot est orthographié قطر dans Badawi et Hinds, 1986, p. 706 et dont l’équivalent fuṣḥā est traditionnellement قطار /qitʕa:r/. Le but est peut-être de marquer la différence avec les articles

دولة قطر 'le Qatar' ou قطر (هندسه) 'diamètre', au prix du choix exclusif de la prononciation du Caire. Ramsis II, le créateur de l'article, a aussi utilisé dans le texte le pluriel avec la suite de graphes 'alif qāf اقْطُرُهُ /ʔutʕura/ 'trains', comme s'il refusait de choisir entre orthographe phonologique et étymologique. L'utilisation des signes diacritiques de vocalisation lui permet de clarifier la prononciation, et donc le sens du mot. Plus loin, il utilise plusieurs mots comprenant un qāf prononcé /ʔ/ au Caire, sans changement orthographique par rapport à la fuṣḥā : قضيب /ʔadʕi:b/ 'rail', ce qui semble indiquer que son utilisation de la 'alif à la place ou en conjonction avec qāf est une expérimentation, dont il n'a pas été trouvé d'autre exemple avec d'autres mots dans cet article.

De même, dans l'article intitulé حقيقه correspondant à l'article "vérité" en français, qāf est remplacé 21 fois par hamza dans حئيثة /ħaʔi:ʔa/ ou حئيي /ħaʔi:ʔi/ 'vrai'. RebeL, le créateur de l'article, écrit pourtant tous les autres qāf étymologiquement comme par exemple يقول /jiʔu:l/ 'il dit'.

Une répartition quasi-systématique du qāf et de la hamza est observable dans l'article L'astronomie chez les Grecs anciens. Son créateur, Masri2wi²³, conserve le qāf quand il est prononcé /q/ en arabe égyptien : اقترح /iqtarah/ 'il suggéra', القرن /ilqarn/ 'le siècle', رقم /raqam/ 'nombre', etc. Il remplace le qāf quand en parler du Caire il est prononcé /ʔ/ par la hamza, écrite souvent selon les règles de la grammaire classique : فرء /farʔ/ 'différence', زء /zaʔʔ/ 'poussa', إياس /ʔija:s/ 'mesure', أبل /ʔabl/ 'avant', وئت /waʔt/ 'temps', منطئة /mantiʔa/ 'région', etc. Il utilise aussi deux hamzas consécutives de manière moins orthodoxe pour les formes de l'élatif : أنصر²⁴ /ʔaʔsʕar/ 'plus court' et أأل /ʔaʔall/ 'moins'. Il lui arrive cependant d'utiliser les deux types d'écritures pour le même mot, sans logique apparente : بيتول /bijʔu:l/ 'il dit' ou d'utiliser deux hamzas différentes pour le même mot : القمر /ilʔamar/ 'la Lune' à la place du classique القمر. De plus, une homographie apparaît avec الأمر /ilʔamr/ 'l'ordre, l'instruction', d'où probablement l'emploi du diacritique fatha sur mīm. Tout cela explique peut-être le choix très majoritaire sur Wikipédia maṣri qui est de conserver qāf, quelque soit sa prononciation effective. Par ailleurs, dans ce même article, un schéma se dessine qui contraste en initiale de mot le "coup de glotte" correspondant à la hamza classique, marqué par une 'alif simple الارض /ilʔardʕ/ 'la Terre' et celui correspondant au qāf classique, marqué par une 'alif surmontée d'une hamza : الأمر /ilʔamar/ 'la Lune', même si Masri2wi utilise aussi wāw en support de hamza à l'initial du mot : أفق /ʔufuq/ 'horizon', à rebours de l'orthographe classique أفق. Cette multitude de réalisations de la hamza complique ainsi considérablement le recours à cette solution pour la représentation de la réalisation de /q/ en /ʔ/ en parler du Caire.

Wāw est aussi utilisé comme support de hamza de manière non orthodoxe dans l'article منبر Minbar²⁵. Ghaly, le rédacteur qui a pourtant contribué à l'élaboration des recommandations d'écriture, a corrigé le verbe يقف 'il se tient debout' en يؤف /juʔaf/ et a ajouté يؤعد /juʔʕud/ à la place du classique يقعد 'il s'assoit'. D'autre part, il a écrit dans plusieurs

²³Masri2wi est un pseudonyme qui signifie "très égyptien", le 2 remplaçant le qāf prononcé /ʔ/ de قوي /ʔawi/ 'très', selon la convention de l'arabizi (Björnsson, 2010, 48 sqq.).

²⁴Cette suite de deux hamzas se trouve en arabe fuṣḥā dans le mot أئمة /ʔaʔimma/ le pluriel de إمام /ʔima:m/ 'imam'.

²⁵plate-forme servant de chaire d'où l'imam fait son sermon lors de la prière du vendredi dans une mosquée.

articles l'expression قوی جداً 'très fort'²⁶ ou قوي للغاية 'extrêmement fort'²⁷, puis Ghalybot, le programme qu'il a conçu (cf. p. 57), transforme للغاية et جداً en اوى, ce qui donne la suite اوى /ʔawi ʔawi/ 'très fort', avec deux mots consécutifs prononcés d'une même façon mais écrits de deux manières différentes, comme envisagé dans l'élaboration des règles (cf. p. 97). Tout ceci illustre la tension existant entre l'idéal d'une écriture phonologique et le pragmatisme de l'écriture étymologique, habituelle.

traduction	API	ق			أ ا			total
			#	%		#	%	
fort, très	/ʔawi/ ²⁸	قوي	35	1,5	أوي	19	1	2.357
					اوي	29	01	
		قوى	1.034	44	أوى	40	1,5	
					اوى	1.200	51	
verre ²⁹	/ʔiza:z/	قزاز	40	18	إزاز	10	5	216
					ازاز	166	77	
combien	/ʔaddi ʔe:/	قد ايه	37	50	اد ايه	37	50	74
autant que	/ʕala ʔadd/	على قد	59	74	على اد	20	25	79
total			1.205	44		1.521	55	2.726

Tab. 4.31: /ʔ/ issu de /q/ transcrit par ق ou أ ا

Le tableau 4.31 montre que pour certains termes, la transcription de /ʔ/ issu de /q/ est majoritairement faite par la 'alif (avec ou sans hamza), même si ce phénomène n'est pas hégémonique. L'influence de Ghalybot peut jouer pour اوى ou ازازه comme il a été vu, mais le phénomène n'est pas négligeable pour d'autres termes. On trouve aussi 24 occurrences de اودام /ʔudda:m/ 'devant' (contre 2.200 de قدام tel qu'orthographié dans Badawi et Hinds (1986, p. 689)). C'est comme si ceux qui écrivent ces mots avec la 'alif ont oublié l'orthographe étymologique, ou tout du moins préfèrent l'ignorer.

4.6 Fricatives dentales /θ/ et /ð/

4.6.1 État des lieux

4.6.1.1 Situation phonologique

En arabe égyptien, on ne prononce pas les consonnes fricatives dentales sourde [θ] et voisée [ð], respectivement transcrites par les graphes ث et ذ en fuṣḥā. Chacune d'entre

²⁶notamment dans يهود امريكان 'Juifs américains' et انجلو ساكسون 'Anglo-Saxons'

²⁷notamment dans فرانكو شكوجور 'Assassinat de Jules César' et فرانكو شكوجور 'Franko Škugor'

²⁹les graphies illustrées peuvent aussi correspondre à /ʔiwi/ "il est devenu fort", /ʔawwa/ "il a renforcé" ou à /quwa/ "force" comme dans العاب قوی "athlétisme" dont 22.199 occurrences figurant dans un titre de paragraphe répété ont été comptées pour une. En dehors de cette répétition, une consultation d'échantillons montre que le sens de "très" ou "fort" est prépondérant.

²⁹Le décompte des occurrences inclut celles de /ʔiza:za/ "bouteille".

elles se prononce de deux manières différentes en devenant fricative alvéolaire ou occlusive alvéolaire, comme l'illustre le tableau 4.32 :

fuṣḥā				‘āmmiyya			traduction	
sourde	[θ]	ث	ثروة	/θarwa/	fricative alvéolaire	[s]	/sarwa/	fortune
			ثلج	/θalʒ/	occlusive alvéolaire	[t]	/talʒ/	glace
voisée	[ð]	ذ	ذكي	/ðakijj/	fricative alvéolaire	[z]	/zaki/	intelligent
			ذراع	/ðiraːʕ/	occlusive alvéolaire	[d]	/diraːʕ/	bras

Tab. 4.32: Devenir de /θ/ et [ð] fuṣḥā en arabe égyptien

Une recension des racines transformées selon un de ces quatre schémas dans *A Dictionary of Egyptian Arabic : Arabic-English* (Badawi et Hinds, 1986) donne une répartition donnant un léger avantage aux fricatives alvéolaires (cf. tableau 4.33)³⁰ :

fuṣḥā	‘āmmiyya	#	%	#	%	#	%
[θ]	[s]	41	56	73	44	166	100
	[t]	32	44				
[ð]	[z]	56	60	93	56		
	[d]	37	40				

Tab. 4.33: Transformation de [θ] et [ð] dans *A Dictionary of Egyptian Arabic*

La consonne issue d'une même racine fuṣḥā peut parfois se prononcer en ‘āmmiyya de deux manières différentes selon le sens mais aussi le schème du mot (cf. tableau 4.34).

fuṣḥā		‘āmmiyya	traduction
ثانية	/θaːnija/	/sanja/	seconde (unité de temps)
		/tanja/	deuxième (féminin)
fuṣḥā		‘āmmiyya	traduction
ذكر	/ðakar(a)/	/zakar/	mentionner
		/dakar/	mâle

Tab. 4.34: Devenir différent de /θ/ et /ð/ pour une même racine

4.6.1.2 Réalisation graphique

En arabe égyptien, le choix du graphe visant à refléter les deux articulations de /θ/ et celles de /ð/ est un enjeu, entre désir de matérialiser la prononciation réelle ou le souhait de garder la transcription de l'arabe littéral.

Rosenbaum (2004, p. 289) explique que lorsque /θ/ se prononce /s/ en ‘āmmiyya, le graphe ث est conservé, comme dans ثواني /sawaːni/ ‘secondes’, phénomène confirmé par la graphie ثانية /sanja/ ‘seconde’ adoptée par Badawi et Hinds (1986, p. 140). Farnawany

³⁰Dans le cas où une même racine a deux transformations possibles, les deux sont comptées séparément.

(1981, p. 82) converge, avec ثورة /sawra/ ‘révolution’. Lorsque /θ/ se prononce /t/ en ‘āmmiyya, le graphe ت a tendance à être adopté, comme dans بعث /baʕat/ ‘il envoya’ (Rosenbaum, 2004, p. 289), orthographié بعث et بعث par Badawi et Hinds (1986, p. 85). Farnawany (1981, p. 51) cite la double graphie اثنا عشر et اثنا عشر pour ʔitna:ʃar/ ‘douze’ ainsi que ثار et ثار pour /ta:r/ ‘vengeance’.

Rosenbaum (2004, p. 290) explique que lorsque /ð/ se prononce /z/ en ‘āmmiyya, le graphe ذ est conservé, comme dans التذاكر /it-taza:kir/ ‘les tickets’, phénomène confirmé par la graphie تذكرة /tazkara/ ‘ticket’ adoptée par Badawi et Hinds (1986, p. 375). Farnawany (1981, p. 90) converge avec ذنب /zamb/ ‘faute, péché’. Lorsque /ð/ se prononce /d/ en ‘āmmiyya, le graphe د a tendance à être adopté, comme dans ما اكذبش عليك /ma ʔakdibʃi ʕale:k/ ‘je ne vais pas te mentir’ (Rosenbaum, 2004, p. 289), verbe dont l’infinitif est orthographié et transcrit كذب et كذب /kidib/, /kadab/ voire /kizib/ ‘mentir’ par Badawi et Hinds (1986, p. 738). Farnawany (1981, p. 52) cite la double graphie عدرا et عدرا pour ʕadra/ ou ʕazra:ʔ/ ‘(la) Vierge’ et الديب /ʔid-di:b/ ‘le loup’ (ibid., p. 60).

Le traitement de la différence phonologique entre fuṣṣḥa et ‘āmmiyya se fait par deux types de choix : le premier respecte l’orthographe étymologique du mot, introduisant une ambiguïté dans le système d’écriture par l’utilisation d’un même graphe pour deux phonèmes distincts. Le deuxième choix correspond à la transcription du changement de phonème, au risque de la création d’éventuels homographes, tels ثعبان pouvant transcrire /taʕba:n/ ‘fatigué’ ou /tiʕba:n/ ‘serpent’ issu de ثعبان (Badawi et Hinds, 1986, p. 129) ou encore بذارة pouvant transcrire /badda:ra/ ‘poudrier’ ou ‘semoir’ issu de بذارة (ibid., p. 57).

4.6.2 Orientations Wikipédia maṣri

4.6.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle ʔarīqa al-kitāba طريقة الكتابة ‘méthode d’écriture’ (Wikipédia, 2025c) comportent les éléments suivants :

ث
لو فيه كلمه من أصل عربي أو أى لغة فيها حرف/نطق "ث" و بيتنطق /س/, يتساب زى ماهو.
و لو بيتنطق /ت/ يتكتب <ت>.

”Si un mot, d’origine arabe ou non, contient la lettre ث qui se prononce /s/, elle reste telle quelle. Si elle se prononce /t/, elle s’écrit ت” (Esperfulmo, le 02/04/2009).

ذ
لو فيه كلمه من أصل عربي أو أى لغة فيها حرف/نطق "ذ" و بيتنطق /ز/, يتساب زى ماهو.
و لو بيتنطق /د/ يتكتب <د>.

”Si un mot, d’origine arabe ou non, contient la lettre ذ qui se prononce /z/, elle reste telle quelle. Si elle se prononce /d/, elle s’écrit د.” (Esperfulmo, le 05/04/2009).

Les concepteurs des recommandations d’écriture sur Wikipédia maṣri explicitent ainsi les usages graphiques déjà repérés.

4.6.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) respecte les recommandations officielles. Il conserve le graphe ث quand il correspond au phonème /s/ en arabe égyptien, ou pour transcrire le phonème /θ/ de mots étrangers. Il le transforme en ت s'il se prononce /t/ (cf. tableau 4.35).

Graphe ث conservé = /s/ ou /θ/			Graphe ث modifié en ت = /t/		
find	replace	traduction	find	replace	traduction
إثيوبيا	اثيوبيا	Éthiopie	كثيراً	كتير	beaucoup
مثلاً	مشل	acteur	ثاني	تاني	deuxième
تأثر	اتأثر	être influencé	ثالث	تالت	troisième
Thatcher	ثاتشر	Thatcher	ثمانينيات	تمانينيات	années 80
تيودور روزفلت	ثيودور روزيفيلت	Theodore Roosevelt	ثعابين	تعابين	serpents

Tab. 4.35: Traitement du graphe ث par Ghalybot

Il conserve le graphe ذ quand il correspond au phonème /z/ en arabe égyptien, ou pour transcrire le phonème /ð/ de mots étrangers. Il le transforme en د s'il se prononce /d/ (cf. tableau 4.36).

Graphe ذ conservé = /z/ ou /ð/			Graphe ذ modifié en د = /d/		
find	replace	traduction	find	replace	traduction
ما إذا	اذا	si	هذا	ده	ce
تم بثها	اتذاعت	a été diffusée	هذه	دى	cette
عدة مرات	كذا مره	plusieurs fois	بذلك	بكه	comme cela
Take That	تيك ذات	Take That	وهكذا	وكده	et comme cela
ريال فاليدوليد	ريال فايادوليد	Real Valladolid	أذرع	دراعات	bras (pluriel)

Tab. 4.36: Traitement du graphe ذ par Ghalybot

4.6.3 Observations dans le corpus

4.6.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-ʿarabiyya

Les graphes ث et surtout ذ font partie de ceux parmi les moins fréquemment utilisés sur Wikipédia maṣri (arz) par rapport à Wikipédia al-ʿarabiyya (ar) (cf. tableau 4.37).

Cela correspond au fait qu'ils transcrivent des phonèmes non prononcés en arabe égyptien. Cependant la relative rareté de ذ par rapport à ث est étonnante car les recommandations sont de conserver ce graphe s'il se prononce [z]. De plus, ث et ذ peuvent être respectivement utilisées pour transcrire les consonnes fricatives dentales sourde [θ] et voisée [ð] de mots étrangers, notamment ceux contenant la séquence <th> en anglais.

graphie	Wikipédia maşri (arz)			Wikipédia al-ʿarabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ث	1.008.905	0,1288%	31	12.334.500	0,4387%	30	12,23
ذ	169.397	0,0216%	38	9.507.807	0,3382%	31	56,13
Corpus	783.599.634	100%	47	2.811.387.118	100%	47	3,59

Tab. 4.37: Fréquence des graphes ث et ذ sur arz et ar

4.6.3.2 Exemples de choix graphiques

Après recensement des occurrences de quelques termes, on observe que la répartition des formes adoptées dans le corpus reflète globalement les recommandations de Wikipédia maşri.

traduction	API	ذ			ز			total
			#	%		#	%	
mémoire	/zakra/	ذاكرة	246	100	ذاكرة	0	0	246
		ذكرة	23	82	ذكرة	5	17	28
torture	/taʕzi:b/	تعذيب	714	99	تعزيب	1	00	715
s'excuser	/(i)ʕtazar/	عتذر	124	99	عتزر	1	00	125
total			1.107	99		7	00	1.114

Tab. 4.38: /z/ transcrit par ذ ou ز

traduction	API	ث			س			total
			#	%		#	%	
richesse	/sarwa/	ثروة	425	99	سروة	1	00	426
investissement	/istisma:r/	استثمار	600	100	استسمار	0	0	600
moderne	/hadi:s/	حديث	13.434	99	حديس	1	00	13.435
exemple	/misa:l/	مثال	505	99	مسال	1	00	506
total			14.964	99		3	00	14.967

Tab. 4.39: /s/ transcrit par ث ou س

Quand le choix d'une orthographe étymologique est préconisé, seules de rares exceptions apparaissent (cf. tableaux 4.38 et 4.39). Quand l'adoption d'une écriture phonémique est préconisée, les cas de respect de l'orthographe étymologique restent nombreux, mais on trouve aussi beaucoup de transcriptions des phonèmes de l'arabe égyptien (cf. tableaux 4.40 et 4.41).

La préférence pour ذ par rapport à د reste prépondérante cependant, alors que ت est globalement plus fréquent que ث. La transformation de أكثر en أكثر par Ghalybot semble contribuer à expliquer la prépondérance de celui-ci. Même chose pour les transformations vers الثاني et الثاني qui inclut les transformations vers الثاني (féminin) et الثانيين (pluriel). Le grand nombre d'occurrences de كثير ne semble pas être le seul résultat des changements

traduction	API	ذ			د			total
			#	%		#	%	
doré	/dahabi/	ذهبي	4.608	95	دهبي	200	04	4.808
		ذهبي	2.425	92	دهبي	196	07	2.621
menton	/daʔn/	ذقن	68	16	دقن	349	83	417
déjeuner	/(i)tyadda/	تغذي	46	90	تغدي	5	09	51
menteur	/kadda:b/	كذاب	134	42	كداب	178	57	312
hérisson	/ʔunfid/	قنفذ	297	93	قنفد	19	06	316
total			7.578	88		947	11	8.525

Tab. 4.40: /d/ transcrit par ذ ou د

traduction	API	ث			ت			total
			#	%		#	%	
plus	/ʔaktar/	أكثر	956	48	أكتر	1.024	51	1980
		اكثر	533	1	اكثر	79.231	99	79.764
le deuxième	/it-ta:ni/	الثاني	5.632	13	التاني	34.806	86	40.438
		التاني	3.611	19	التاني	14.614	80	18225
trois	/tala:t/	ثلاث	5.132	45	تلات	6.068	54	11.200
huit	/tama:n/	ثمان	10.512	57	ثمان	7.772	42	18.284
beaucoup	/kiti:r/	كثير	2.882	19	كتير	11.660	80	14.542
renard	/taʕlab/	ثعلب	601	90	تعلب	64	09	665
lourd	/tiʔi:l/	ثقيل	338	65	تقيل	177	34	515
total			30.197	16		155.416	83	185.613

Tab. 4.41: /t/ transcrit par ث et ت

opérés par Ghalybot, qui ne remplace que كَثِيرًا, mais pas les autres formes du mot sans la fathātān. Cette répartition reflète donc probablement des habitudes d'écriture déjà ancrées dans la population. On observe ainsi une hésitation entre la fidélité à l'orthographe fuṣḥā et l'émancipation vers une transcription phonémique de la ʿāmmiyya.

4.7 Fricatives dentale /ð^ʕ/, alvéolaire /z^ʕ/ et occlusive alvéolaire /d^ʕ/ pharyngalisées

4.7.1 État des lieux

4.7.1.1 Situation phonologique

Selon Gadalla (2000, p. 17) , [ð^ʕ], la version pharyngalisée de la consonne fricative dentale voisée [ð], transcrite par le graphe <ظ> en «arabe classique», devient la fricative alvéolaire pharyngalisée [z^ʕ] en «arabe standard» et en arabe égyptien, où elle devient, pour certains mots l'occlusive alvéolaire pharyngalisée [d^ʕ] (cf. tableau 4.42).

"arabe classique"				"arabe standard"		arabe égyptien		traduction
ظ	ظَهْر	[ðʕ]	/ðʕuhr/	[zʕ]	/zʕuhr/	[dʕ]	/dʕuhr/	midi
	ظَهَرَ		/ðʕahara/		/zʕahara/	[zʕ]	/zʕahar/	apparaître

Tab. 4.42: Passage de /ð^ʕ/ en fuṣḥā à /z^ʕ/ ou /d^ʕ/ en ‘āmmiyya

Ainsi, concernant le glissement de /ð^ʕ/ vers /d^ʕ/ (ou /z^ʕ/), dans *A Dictionary of Egyptian Arabic : Arabic-English*, on trouve la prononciation de /d^ʕufr/ ou /d^ʕifr/ ‘ongle’, /z^ʕill/ ou /d^ʕill/ ‘ombre’ (Badawi et Hinds, 1986, p. 523), /d^ʕalma/ ‘obscurité’ (ibid., p. 524), /d^ʕahr/ ‘dos’ ou /d^ʕuhr/ ‘midi’ (ibid., p. 525-6), /ʕad^ʕm/ ‘os’ (ibid., p. 583), /nid^ʕif/ ou /nud^ʕuf/ ‘devenir propre’ (ibid., p. 867), ainsi que pour certains de leurs dérivés. De même pour /ħad^ʕi:r/ ‘couloir’ (ibid., p. 212), /maħfaz^ʕa/ et /maħfad^ʕa/ ‘portefeuille’ (ibid., p. 213). On note aussi /maħd^ʕijja/ ‘femme préférée’ (ibid., p. 211), /fad^ʕja/ ou /fat^ʕja/ ‘écharde’ (ibid., p. 464), /nad^ʕd^ʕa:ra/ ‘lunettes’ (ibid., p. 870), /hufa:d^ʕ/ ‘couche-culotte’ (ibid., p. 213).

On observe aussi une sorte de mouvement inverse dans lequel le phonème /d^ʕ/ de la fuṣḥā devient /z^ʕ/ voire /z/ en ‘āmmiyya. Badawi et Hinds (ibid., p. 555), donnent la prononciation ‘āmmiyya pour /z^ʕarat^ʕ/ ‘péter’ et /z^ʕabat^ʕ/ ou /zabat^ʕ/ ‘ajuster’. Pour ce dernier, et ses dérivés, ce dictionnaire donne les prononciations ‘āmmiyya rassemblées dans le tableau 4.43.

fuṣḥā	/d ^ʕ abat ^ʕ a/		/d ^ʕ abt ^ʕ /			/d ^ʕ a:bit ^ʕ /		
	[z ^ʕ]	[z]	[d ^ʕ]	[z ^ʕ]	[z]	[d ^ʕ]	[z ^ʕ]	[z]
‘āmm	/z ^ʕ abat ^ʕ /	/zabat ^ʕ /	/d ^ʕ abt ^ʕ /	/z ^ʕ abt ^ʕ /	/zabt ^ʕ /	/d ^ʕ a:bit ^ʕ /	/z ^ʕ a:bit ^ʕ /	/za:bit ^ʕ /
yya	/z ^ʕ abat ^ʕ il- niz ^ʕ a:m/	/zabat ^ʕ ilhisa:b/	/d ^ʕ abt ^ʕ in- nafs/	/biz ^ʕ z ^ʕ abt ^ʕ /	/bizzabt ^ʕ /			
traduction	rétablir l'ordre	mettre de l'ordre dans le compte	sang-froid	à [heure] précisément		officier		

Tab. 4.43: Passage de /d^ʕ/ en fuṣḥā à /z^ʕ/ en ‘āmmiyya

Outre des formes dérivées du verbe, sont aussi cités /maz^ʕbu:t^ʕ/ et /mazbu:t^ʕ/ ‘précis, exact’, ainsi que le classicisme : /ʔind^ʕiba:t^ʕ/ ‘discipline’, sans alternative en [z^ʕ] ou [z].

4.7.1.2 Réalisation graphique

Pour ce qui est du glissement de /ð^ʕ/ vers /d^ʕ/, dans *A Dictionary of Egyptian Arabic : Arabic-English* (ibid.), on trouve les graphies fuṣḥā et ‘āmmiyya et la prononciation ‘āmmiyya pour ظفر et ضرر /d^ʕufr/ ou /d^ʕifr/ ‘ongle’, ظل et ضل /z^ʕill/ ou /d^ʕill/ ‘ombre’ (ibid., p. 523), ظلمة et ضلمة /d^ʕalma/ ‘obscurité’ (ibid., p. 524), ظهر et ضرر /d^ʕahr/ ‘dos’ ou /d^ʕuhr/ ‘midi’ (ibid., p. 525-6), عظم et عضم /ʕad^ʕm/ ‘os’ (ibid., p. 583), نظيف et نظيف /nid^ʕif/ ou /nud^ʕuf/ ‘devenir propre’ (ibid., p. 867), ainsi que pour certains de leurs dérivés. De même pour حظير et حضير /ħad^ʕi:r/ ‘couloir’ (ibid., p. 212), محفظة et محفظة /maħfaz^ʕa/ et

/maḥfad^ʕa/ ‘portefeuille’ (Badawi et Hinds, 1986, p. 213). On note aussi le contraste entre l’écriture et la prononciation de محظية /maḥd^ʕijja/ ‘femme préférée’ (ibid., p. 211), شظية /ʃad^ʕja/ ou /ʃat^ʕja/ ‘écharde’ (ibid., p. 464), نظارة /nad^ʕd^ʕa:ra/ ‘lunettes’ (ibid., p. 870), حفاظ /ḥufa:d^ʕ/ ‘couche-culotte’ (ibid., p. 213).

Pour le mouvement inverse dans lequel le phonème /d^ʕ/ de la fuṣṣḥa devient /z^ʕ/ voire /z/ en ‘āmmiyya. Badawi et Hinds (ibid., p. 555), donnent les graphies fuṣṣḥa et ‘āmmiyya et la prononciation ‘āmmiyya pour ضراط /z^ʕarat^ʕ/ ‘péter’ et ضبط /z^ʕabat^ʕ/ ou /zabat^ʕ/ ‘ajuster’. Dans *A Dictionary of Egyptian Arabic : Arabic-English* (ibid.), chacun de ces mots est orthographié en graphie arabe à la fois avec ض et avec ظ, mais jamais avec ز.

Rosenbaum (2004, p. 290) explique que /d^ʕuhr/ ‘midi’ est parfois écrit selon l’orthographe fuṣṣḥa ظهر, parfois selon la prononciation ‘āmmiyya ضهر, tandis que /d^ʕahr/ ‘dos’ est occasionnellement écrit selon l’orthographe fuṣṣḥa ظهر, mais généralement selon la prononciation ‘āmmiyya ضهر. Farnawany (1981, p. 102) cite uniquement l’orthographe fuṣṣḥa de مضبوط /maz^ʕbu:t^ʕ/ ou /mazbu:t^ʕ/ ‘exactement’ et ضابط /z^ʕa:bit^ʕ/ ou /za:bit^ʕ/ ‘officier’, puis la graphie reflétant la prononciation de بالضبط /biz^ʕ-z^ʕabt^ʕ/ ou /biz-zabt^ʕ/ (ibid., p. 106).

4.7.2 Orientations Wikipédia maṣri

4.7.2.1 Recommandations officielles

Pour ظ ou ض, aucune recommandation n’est faite, mais sur la page de discussion (archive 1) ramsis s’interroge le 10/05/2008 sur la graphie à adopter entre المظلومين ou المزلومين pour ‘les opprimés’.

4.7.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) ne traite pas le passage de ظ en ض. Par contre, il transforme ض en ظ dans les deux cas recensés dans le tableau 4.44.

find	replace	traduction	find	replace	traduction
ضباط	ظباط	officiers	ضبط استنادي	ظبط استنادي	contrôle d’autorité

Tab. 4.44: Traitement du graphe ظ par Ghalybot

4.7.3 Observations dans le corpus

4.7.3.1 Fréquence d’utilisation des graphes par rapport à Wikipédia al-‘arabiyya

Le graphe ظ est légèrement moins fréquent sur Wikipédia maṣri (arz) par rapport à Wikipédia al-‘arabiyya (ar) (cf. tableau 4.45), probable résultat de l’équilibre entre la tendance à rendre phonétiquement le remplacement en arabe égyptien de la fricative alvéolaire pharyngalisée [z^ʕ] par l’occlusive alvéolaire pharyngalisée [d^ʕ] par exemple pour ضهر /d^ʕahr/ ‘dos’ et la tendance phonétique inverse pour ظابط /z^ʕa:bit^ʕ/ ‘officier’ (cf. p. 105).

graphie	Wikipédia mašri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ظ	972.875	0,1242%	32	4.427.141	0,1575%	34	4,55
ض	3.299.101	0,4210%	26	13.122.675	0,4668%	28	3,98
Corpus	783.599.634	100%	47	2.811.387.118	100%	47	3,59

Tab. 4.45: Fréquence des graphes ظ et ض sur arz et ar

4.7.3.2 Exemples de choix graphiques

La recension d’occurrences dans une sélection de mots indique une forte tendance à conserver l’orthographe étymologique.

C’est notamment le cas pour le passage vers ض, pour lequel presque chaque terme observé est écrit majoritairement avec le graphe ظ, soit selon l’orthographe étymologique, comme observé dans le tableau 4.46.

traduction	API	ظ			ض			total
			#	%		#	%	
obscurité	/dˤalma/	ظلمة	176	79	ضلمة	45	20	221
		ظلمه	110	70	ضلمه	46	29	156
obscurité	/dˤala:m/	ظلام	721	98	ضلام	11	01	732
l’ombre	/il-dˤill/ /il-zˤill/	الظل	1.157	98	الضل	17	01	1.174
le dos / midi	/il-dˤahr/ /il-dˤuhr/	الظهر	596	80	الضهر	149	20	745
lunettes	/nadˤdˤa:ra/	نظارة	218	95	نضارة	10	04	228
		نظاره	137	93	نضاره	10	06	147
lunettes	/nadˤdˤara:t/	نظارات	96	96	نضارات	3	03	99
propre	/nidˤi:f/	نظيف	249	80	نضيف	60	19	309
propreté	/nadˤa:fa/	نظافة	40	97	نضافة	1	02	41
		نظافه	9	47	نضافه	10	52	19
os (pluriel)	/ʕidˤa:m/ /ʕizˤa:m/	عظام	367	97	عضام	10	02	377
total			3.876	91		372	08	4.248

Tab. 4.46: /dˤ/ transcrit par ظ ou ض

Un échantillonnage de 20 occurrences pour chacun des termes الظهر et الظهر montre que 60% des premiers ont le sens de ‘midi’, et que 60% des seconds ont le sens de ‘dos’, ce qui pour ce dernier confirme l’appréciation de Rosenbaum (2004, p. 290) quand il écrit que celui-ci ”est généralement écrit comme il est prononcé en ‘āmmyya”³¹. Par ailleurs, un article intitulé حفاضه traite de la ‘couche-culotte’ avec transcription de la prononciation ‘āmmyya du terme apparenté à la racine fuṣḥā ح ف ظ.

Pour le passage vers ظ, voire ز, le constat est plus nuancé. Pour /zˤa:bitˤ/ et /zˤubba:tˤ/, dont Badawi et Hinds (1986, p. 555) donnent la prononciation alternative fuṣḥā /dˤa:bitˤ/

³¹ “[It] is generally written as it is pronounced in ‘āmmyya.”

/d^ʕubba:t^ʕ/, c'est l'orthographe fuṣṣhā qui reste majoritaire³², mais pour /biz^ʕ-z^ʕabt^ʕ/ et /maz^ʕbu:t^ʕ/, bien moins représentés, c'est l'inverse, et on trouve même des occurrences de recours au graphe ʒ correspondant aux prononciations /zabt^ʕ/ et /mazbu:t^ʕ/ aussi indiquées dans Badawi et Hinds (1986) (cf. tableau 4.47).

traduction	API	ض			ظ			ز			total
			#	%		#	%		#	%	
officier	/z ^ʕ a:bit ^ʕ /	ضابط	15.887	76	ظابط	4.825	23	زابط	1	0	20.713
officiers	/z ^ʕ ubba:t ^ʕ /	ضباط	2.238	63	ظباط	1.273	36	زباط	0	0	3.511
exactement	/biz ^ʕ -z ^ʕ abt ^ʕ /	بالضبط	31	08	بالظبط	334	88	بالزبط	11	02	376
exact	/maz ^ʕ bu:t ^ʕ /	مضبوط	19	25	مظبوط	46	62	مزبوط	9	12	74
total			18.175	73		6.478	26		21	00	24.674

Tab. 4.47: /z^ʕ/ transcrit par ض, ظ ou ز

La transcription de la prononciation 'āmmiyya apparaît minoritaire dans les occurrences observées, même quand Ghalybot transforme ضباط en ظباط. L'influence de l'orthographe étymologique reste prépondérante.

4.8 /g/ consonne occlusive vélaire voisée

4.8.1 État des lieux

4.8.1.1 Situation phonologique

Une des particularités de l'arabe égyptien parlé au Caire (et donc la variété la plus prestigieuse) est le recours à la consonne occlusive vélaire voisée /g/ pour la prononciation du phonème /dʒ/ correspondant en fuṣṣhā. Le mot /dʒama:l/ est ainsi prononcé /gama:l/ au Caire, comme pour le prénom de Ġamāl 'Abd al-Nāṣir, le célèbre homme d'État égyptien.

On trouve aussi ce phonème dans la prononciation des mots d'origine étrangère, comme dans /ʔingili:zi/ 'anglais'.

Mais inversement on trouve aussi le phonème /z/ de certains mots étrangers transformés en /g/ en parler du Caire : /gurna:l/ 'journal' (ibid., p. 157), /guɣrafja/ 'géographie' (ibid., p. 163) et /biɣa:ma/ ou /biga:ma/ (ce dernier noté comme relevant du niveau bas du registre de langue) pour 'pyjama' dans Badawi et Hinds (ibid., p. 53) (cf. aussi Farnawany (1981, p. 83)).

Parallèlement à la prononciation /g/ du Caire et de la partie centrale du Delta, la prononciation /dʒ/ s'entend dans le régiolecte ṣa'īdi (de Haute Égypte) (Khalafallah, 1969, p. 41) et dans la partie est du Delta (Manfred Woidich, 2008, p. 6).

³²L'occurrence avec ز vient du vers *أمة و ما لهاش لا زابط ولا رابط* "une nation et elle n'a ni officier, ni défenseur" du poème *عبد الناصر يا عم* *الله يجازيك يا عم* "Que Dieu te récompense, ô Oncle Abdel Nasser" du poète 'Abd ir-Rahmān il-Abnūdī.

4.8.1.2 Réalisation graphique

La consonne occlusive vélaire voisée /g/ du Caire est généralement transcrite par le graphe ġīm ج de l'arabe littéral.³³ Ainsi, le mot /gama:l/ 'beauté' s'y écrit جمال.

Le lecteur garde ainsi la liberté d'interpréter la valeur phonétique de ce graphe, et donc de lire حجاج "pèlerins" comme se prononçant /ħudʒdʒa:dʒ/ (fuṣṣhā ou parler ʕa'īdi) ou /ħugga:g/ (parler du Caire).

Pour les mots d'origine étrangère, c'est au rédacteur de faire un choix entre différents graphes pour transcrire le phonème /g/. Pour un locuteur cairote, ce sera naturellement le graphe ġīm ج. C'est le cas du mot /ʔingili:zi/ 'anglais', qui est généralement transcrit en Égypte par إنجليزي (Wehr et Cowan, 1979, p. 38) à la différence d'autres zones du monde arabe qui transcrivent إنكليزي avec le graphe kāf ك (ibid., p. 40)³⁴.

Traditionnellement, l'arabe classique transcrit l'occlusive vélaire voisée /g/ des noms étrangers avec le graphe ġayn غ qui correspond habituellement à la fricative vélaire voisée /ɣ/ considérée comme voisine. D'où l'emploi de غرناطة /ɣarna:tʔa/ pour le nom de la ville de Grenade, notamment sur Wikipédia al-'arabiyya³⁵.

4.8.2 Orientations Wikipédia maṣri

4.8.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle Ṭarīqa al-kitāba طريقة الكتابة 'méthode d'écriture' (Wikipédia, 2025c) comportent l'élément suivant :

ج
في مصر، الاسامي الاجنبية الى فيها نطق [g] بتكتب بالحرف العربي ج، مش بالغ.
"En Égypte, les noms étrangers qui ont la prononciation [g] s'écrivent avec la lettre arabe ج et pas avec la lettre غ." (Esperfulmo, le 02/04/2009).

Le graphe ج est donc conseillée pour la transcription de /g/ des noms étrangers comme celui de la ville espagnole "Grenade", traditionnellement transcrit غرناطة en arabe littéral et dont l'article sur Wikipédia maṣri porte le titre جرانادا correspondant à la prononciation espagnole de ce nom.

³³Le graphe ġīm ج se caractérise par la multiplicité de ses réalisations phonémiques dans le monde arabe. Il correspond à [dʒ] dans la péninsule arabique, dans certaines parties du Levant, en Irak et dans le nord de l'Algérie, prononciation également considérée comme standard pour l'arabe littéral, [ʒ] dans la majeure partie de l'Afrique du Nord-Ouest et des villes du Levant, [g] au Caire, centre et nord-est du delta du Nil, sur la côte du Yémen et sur la côte sud d'Oman, ainsi que [j] au Soudan (Youssef, 2021, p. 7).

³⁴Dans le monde arabe, pour transcrire /g/, sont aussi utilisées les variantes suivantes : گ avec un trait dessus pour écrire les mots dialectaux en arabe du Golfe et de Mésopotamie (e.g. /gəmət/ "lune"), گ with trois points pour Maroc (e.g. /ʔaga:di:r/ "Agadir"), في avec trois points en Algérie et en Tunisie (e.g. /ga:bis/ "Gabès"). Paradoxalement, le titre de *A Dictionary of Egyptian Arabic : Arabic-English* est traduit en arabe avec la graphie إنكليزي (Badawi et Hinds, 1986, p. V). Ceci est probablement lié au fait que ce dictionnaire a été publié au Liban.

³⁵On trouve aussi sur Wikipédia al-'arabiyya l'article intitulé جوجل pour "Google", également orthographié جوجل et جوجل dans la phrase d'introduction de l'article, ce qui semble indiquer qu'aucun choix ne semble s'imposer dans l'usage fuṣṣhā moderne pour transcrire /g/. L'article correspondant sur Wikipédia en arabe marocain est intitulé جوجول et en arabe égyptien جوجول.

4.8.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) remplace le graphe غ par ج dans environ 25 mots dont غيتار 'guitar', ديموغراف 'démographe', فيلموغرافيا 'filmographie', غواتيمالا 'Guatemala', هونغ كونغ 'Hong Kong', ريغان 'Reagan', هابسبورغ 'Habsbourg', هاستينغز 'Hastings'. Il transforme غلاسكو en جلاسيكو pour 'Glasgow'. Il applique ainsi les recommandations officielles pour ces mots.

4.8.3 Observations dans le corpus

4.8.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-'arabiyya

Le graphe غ est bien moins fréquent sur Wikipédia maşri (arz) par rapport à Wikipédia al-'arabiyya (ar) alors que le graphe ج est bien plus fréquent, comme le montre le tableau 4.48.

graphe	Wikipédia maşri (arz)			Wikipédia al-'arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
غ	3.226.804	0,4118%	27	15.515.668	0,5519%	27	4,81
ج	15.821.354	2,0191%	18	36.108.823	1,2844%	21	2,28
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.48: Fréquence des graphes غ et ج sur arz et ar

Ces différences tendent à confirmer le choix fait d'utiliser le graphe ج pour transcrire le phonème /g/ des mots étrangers ou d'origine étrangère à la place du graphe غ.

4.8.3.2 Exemples de choix graphiques

On observe que les rédacteurs sur Wikipédia maşri choisissent naturellement le graphe گīm ج pour transcrire /g/, notamment dans les mots d'origine étrangère, comme pour انگليز 'anglais', dont on trouve 150.964 occurrences dans le corpus, contre 236 pour انكليز, qui utilise le graphe kāf ك et une seule occurrence de باللغة الإنكليزية 'en langue anglaise' avec غ (Ghassan Rahbani).

Cependant, l'alternance la plus courante est entre le graphe گīm ج préféré au Caire et le graphe ḡayn غ favorisé en fuṣḥā. Dans le tableau 4.49, sont présentés différents exemples d'utilisation de ces deux graphes pour un même mot.

Le choix très majoritaire est celui du graphe گīm ج pour transcrire le phonème /g/ des mots d'origine étrangère, qu'ils soient traités par Ghalybot, comme جواتيمالا 'Guatemala' ou non comme pour أنجولا 'Angola'. Le graphe ḡayn غ est majoritairement conservé pour 'Afghanistan' /ʔafɣanista:n/ (dont la prononciation avec /ɣ/ est celle indiquée par Badawi et Hinds (1986, p. 27)) et pour 'gorille' /ɣurilla/ (aussi transcrit avec /ɣ/ par Badawi et Hinds (ibid., p. 621)). Pour ce dernier, près d'un tiers des occurrences utilisent cependant le graphe گīm ج, peut-être par influence de la prononciation anglaise. Pour la transcription des terminaisons -graph(i)e, ḡayn غ est préféré pour ديموغراف 'démographe', malgré le remplacement par گīm ج effectué par Ghalybot. Par contre, ce dernier graphe est le seul

traduction	API	ج			غ			total
			#	%		#	%	
Guatemala	/gwatimala/	جواتيمالا	2.252	99	غواتيمالا	39	01	2.291
Angola	/ʔanguːla/	أنجولا	2.620	95	أنغولا	131	05	2.751
guitare	/gitaːr/	جيتار	14.293	99	غيتار	37	1	14.330
gorille	/ɣurilla/	جورिला	33	33	غورिला	66	67	99
Afghanistan	/ʔafɣanistaːn/	أفغانستان	463	14	أفغانستان	2.660	86	3.123
démographe	/dimugraːf/	ديموغراف	644	30	ديموغراف	1.559	70	2.203
pornographie	/burnugrafja/	بورنوجرافيا	8.223	100	بورنوغرافيا	0	0	8.223
total			28.528	86		4.492	14	33.020

Tab. 4.49: /g/ ou /ɣ/ transcrits par ج ou غ

utilisé pour بورنوجرافيا ‘pornographie’ (sans intervention de Ghalybot), peut-être ici aussi par influence de l’anglais.

4.9 [ʒ], [v] et [p] phones allogènes

4.9.1 État des lieux

4.9.1.1 Situation phonologique

Les phones [ʒ], [v] et [p] peuvent être considérés comme des phones allogènes en arabe du Caire en ce qu’ils s’entendent essentiellement dans des mots d’origine étrangère relativement récente. Ils ne font pas partie des consonnes entrant dans la formation de racines trilitères ou quadrilitères des mots arabes fuṣḥā ou intégrés à l’arabe du Caire.

4.9.1.1.1 [ʒ]

Si la consonne affriquée palato-alvéolaire voisée /dʒ/ de la fuṣḥā est remplacée par l’occlusive vélaire voisée /g/ en parler du Caire, on y entend aussi la fricative palato-alvéolaire voisée [ʒ]. Elle apparaît en alternative à /g/ en face du graphe ج dans le tableau des consonnes du dictionnaire de Badawi et Hinds (1986, p. XVII).

Dans Badawi et Hinds (ibid.), [ʒ] apparaît marginalement avant tout pour des mots d’origine étrangère comme /ʒaːd/ ‘jade’ (ibid., p. 146), /ʒirs/ ou /ʒirseː/ ‘jersey, pullover’ (ibid., p. 155), /garaːʒ/ ou /garaːf/ ‘garage’ (ibid., p. 156), /ʒiruː/ ‘emmanchure’ de l’italien *giromanica* (ibid., p. 157), /ʒakitta/ ‘veste’ de l’italien *giacchetta*, /ʒilaːti/ ‘glace’ de l’italien *gelato* (ibid., p. 164), /ʒili/ ‘gelée, dessert’ de l’anglais *jelly* (ibid., p. 164), /ʒinz/ ‘pantalon jeans’, /ʒant/ ‘jante’ (ibid., p. 176), /ʒiːb/ ‘Jeep’, /ʒiːba/ ‘jupe’, /ʒijuluʒja/ ‘géologie’ (ibid., p. 185). Grâce à /ʒiːb/ ‘Jeep’, on peut former la paire minimale avec /giːb/ ‘apporte’ (ibid., p. 184-185), ce qui peut permettre de considérer [ʒ] comme un phonème de l’arabe du Caire.

4.9.1.1.2 [v]

À côté de la consonne fricative labio-dentale sourde /f/, son équivalente voisée /v/ est considérée comme difficile à prononcer pour les locuteurs arabes, mais moins que /p/ (Al-Zoubi, 2019). Sa transcription apparaît en alternative à /f/ en face du graphe ف dans le tableau des consonnes du dictionnaire de Badawi et Hinds (1986, p. XVII). On trouve dans ce dernier des transcriptions telles que : /vitami:n/ ou /fitami:n/ ‘vitamine’ (ibid., p. 640), /vazli:n/ ou /fazli:n/ ‘Vaseline’ (ibid., p. 655), /villa/ ou /filla/ ‘villa ou duplex’ (ibid., p. 671), /vanilja/ ou /fanilja/ ou même /banilja/ ‘vanille’ (ibid., p. 673), /vuwa:l/ ou /fuwa:l/ ou même /buwa:l/ ‘voile coton’ (ibid., p. 675), /vi:za/ ou /fi:sa/ ‘visa’ (ibid., p. 679), etc.

Grâce à /villa/ ‘villa ou duplex’, on peut former la paire minimale avec /filla/ ‘un bouchon de liège’ (ibid., p. 671), ce qui oriente vers la prise en compte de [v] comme un phonème de l’arabe du Caire. Cependant, cette prononciation en [v] est généralement présentée avec une alternative en /f/ ou en /b/ dans un niveau de langue plus populaire³⁶ où [v] ne semble pas pouvoir être compté comme phonème à part entière.

4.9.1.1.3 [p]

Face à la consonne occlusive bilabiale voisée /b/, son équivalente sourde [p] est naturellement très difficile à entendre et à prononcer pour les locuteurs arabes (Al-Zoubi, 2019). Elle est absente du tableau des phonèmes de l’arabe et de la plupart de ses dialectes. Elle n’apparaît pas en alternative à /b/ en face du graphe ب dans le tableau des consonnes du dictionnaire de Badawi et Hinds (1986, p. XVII). On trouve pourtant un certains nombres d’emprunts de langues étrangères comportant à l’origine ce son : /ba:fə/ ‘Pasha, Monsieur’, du turc *paşa* (ibid., p. 49), /bila:ʒ/ ‘plage’ du français (ibid., p. 95), /babu:r/ ‘réchaud à gaz’ de l’italien *vapore* (ibid., p. 50), /baraʃut/ ‘parachute’ du français (ibid., p. 65), /buli:s/ ‘police’ du français (ibid., p. 103), /bantʰalo:n/ ‘pantalon’ du français (ibid., p. 107), /kubri/ ‘pont’ du turc *köprü* (ibid., p. 732), /kabtin/ ‘capitaine’ de l’anglais *captain* (ibid., p. 730), /bu:r/ ‘Port’ comme dans /bu:r saʕi:d/ ‘Port Said’ (ibid., p. 111), etc. Dans *A Dictionary of Egyptian Arabic : Arabic-English*, il n’a été trouvé qu’une transcription indiquant la prononciation /p/ pour un de ces termes, à savoir /patina:ʒ/ en parallèle avec /batina:ʒ/ ‘patinage’ (ibid., p. 53).

4.9.1.2 Réalisation graphique

Pour transcrire les phones allogènes à l’arabe, les graphes arabes des phonèmes proches sont généralement employés, mais d’autres issus de l’alphabet perso-arabe³⁷ sont parfois utilisés.

4.9.1.2.1 ج et چ

A Dictionary of Egyptian Arabic : Arabic-English (ibid., p. 146-186), transcrit chaque entrée comportant /g/ ou [ʒ] dans sa prononciation par le graphe ج. Farnawany (1981,

³⁶ ‘āmmyyatu l-‘ummiyyīn (dialecte des illettrés) de Badawi (1973)

³⁷ Cinq graphes ont été créés spécifiquement dans l’alphabet perso-arabe pour représenter les sons persans inconnus en arabe littéral, à savoir پ pour /p/, ژ pour /ʒ/, چ pour /tʃ/, گ pour /g/ et ف pour /β/.

p. 83) confirme avec l'exemple جاكته /ʒakitta/ 'veste', mais pointe aussi l'emploi du graphe ج³⁸ avec چاكته. Rosenbaum (2004, p. 288) cite à la fois جورج et چورچ /ʒorʒ/ pour 'George'. Il pointe l'utilisation simultanée des deux graphes dans un même mot pour transcrire les deux sons, comme dans جراج /gara:ʒ/ 'garage, parking', comme on peut le voir dans les rues du Caire ou dans le titre sur la couverture du roman de Ḥammād 'Ilīwa *Garāğ al-Zamālik* الزمالك جراج 'Le garage de Zamalik' publié en 2023 par Dār al-Riwāq. Une paire minimale graphique peut être observée entre جيم /gi:m/ 'la lettre ġīm' (Badawi et Hinds, 1986, p. 186) et چيم /ʒi:m/ 'salle de sport' (Aldrich, 2020, p. 139), ou encore entre جاز /ga:z/ 'paraffine, kérosène' (Badawi et Hinds, 1986, p. 146) et الجاز /iʒ-ʒa:z/ 'le jazz' (Aldrich, 2020, p. 132). Avec ces emplois, ج s'approche du statut de graphème en arabe égyptien.

4.9.1.2.2 ف et ث

A Dictionary of Egyptian Arabic : Arabic-English (Badawi et Hinds, 1986, p. 637-681), transcrit chaque entrée commençant par /f/ ou [v] dans sa prononciation avec le graphe ف. Farnawany (1981, p. 108) documente le même usage avec فازه /va:za/ 'vase'. Il indique que, très rarement, est utilisé le graphe ث³⁹ comme pour فيله /villa/ 'villa'. Rosenbaum (2004, p. 289) cite بروقة للجريمة /biru:va lilgari:ma/ 'répétition du crime' dans lequel on trouve le graphe ث sur la couverture et sur la page de titre, mais pas dans le reste du texte de cette pièce de 1988 de Muḥammad Sālim. Le graphe ث est aussi utilisé, notamment pour transcrire des noms de marques comme Vodafone Égypte فودافون مصر "Vodafone Maṣr", le plus grand opérateur de réseau mobile en Égypte.

4.9.1.2.3 ب et پ

Dans *A Dictionary of Egyptian Arabic : Arabic-English*, même pour /patina:ʒ/ 'patinage' (Badawi et Hinds, 1986, p. 53), la transcription du son /p/ est faite par le graphe بā' ب utilisé dans ce dictionnaire pour transcrire le son /b/. Le graphe پ (bā' à trois points)⁴⁰ n'est pas non plus mentionné par Rosenbaum (2004). Farnawany (1981, p. 176) y fait référence pour indiquer : "La lettre - پ - n'a pas été rencontrée dans les textes."⁴¹ et il précise : "Le son [p], en revanche, n'est jamais représenté par - پ -, bien que ce symbole

³⁸Le graphe ج dérivé de ġīm ج par l'ajout de deux points est utilisé dans l'alphabet perso-arabe, pour transcrire /ġ/. Il est utilisé en arabe irakien avec la même valeur, en arabe levantin pour transcrire /g/ et en arabe du Caire pour transcrire /ʒ/. Il y est appelé ġīm bi talat no'at جيم بثلاث نقط "ġīm à trois points". Dans l'alphabet perso-arabe (entre autres), /ʒ/ est transcrit par ج.

³⁹ث est un graphe dérivé de ف par l'ajout de deux points et est à l'origine l'un des graphes ajoutés à l'alphabet perso-arabe pour écrire le persan moderne, où il transcrivait /β/. Ce graphe n'est plus utilisé en persan, /β/ étant devenu /b/. ث transcrit actuellement /v/ dans les alphabets kurde et swahili, notamment, ce phonème étant transcrit par و dans l'alphabet perso-arabe ou un de ses dérivés dans des langues apparentées. /v/ est transcrit par و en Algérie et en Tunisie pour les noms étrangers et les mots d'emprunt, de manière à le différencier de ق (qāf à trois points) qui y transcrit /g/.

⁴⁰Le graphe پ (bā' à trois points) est notamment utilisé pour transcrire /p/ en persan, kurde, pachtoun, baloutche et dans certaines langues turques comme le ouïghour.

⁴¹"Der Buchstabe - پ - wurde in den Texten nicht angetroffen."

soit parfois utilisé dans les journaux.”⁴². Aldrich (2020, p. 132) l'utilise pour transcrire environ 17 occurrences du son [p] dans des mots d'origine étrangère, dont پیتزا 'pizza'.

Le graphe پ est aussi occasionnellement utilisé en Égypte, lors de la translittération de noms étrangers et d'emprunt à la place de bā' ب. On le lit sur une plaque de la rue Champollion شارع شامبوليون à Alexandrie, et sur la devanture de Francis Papazian فرنسيس بابازيان, un horloger historique d'origine arménienne au Caire, dont d'anciennes publicités montrent que son nom était déjà transcrit avec ce graphe au cours du xx^e siècle. On trouve aussi, sur <https://www.marefa.org/>, une encyclopédie en ligne en arabe fuṣḥā similaire à Wikipédia, un article intitulé ب پ (BP)⁴³ au sujet de la compagnie pétrolière britannique et qui différencie par la graphie la version voisée de l'occlusive bilabiale de sa version sourde.

4.9.2 Orientations Wikipédia maṣri

4.9.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle Ṭarīqa al-kitāba طريقة الكتابة 'méthode d'écriture' (Wikipédia, 2025c) donnent les instructions suivantes :

يمكن تستعمل حروف بتوَّحّ النطق جَوَّ الصَّفح, زى پ چ ف.
السبب هو ان الحروف دى فى العناوين بتخلّى المقالات ماتظهرش فى سايئات التدوير, زى مثلاً جوجل, لإن ما حدش هايكتب بيهم و هو بيدور على جوجل.
حرف چ هو لنطق /t/ زى ما يستخدم فى مصر, مش /t/ زى الفارسى! النطق التانى بيتكتب بـ تش.
حرف ف هو لنطق /v/ زى ما يستخدم فى مصر, مش فى التونسى-الجزائرى الى يستخدم لنطق /f/. النطق التانى بيتكتب بحرف الج.
اكتب فيفا او فيفا. اكتب ايفا او ايفا, مافيش مشكله.

”Tu peux utiliser des lettres qui clarifient la prononciation dans le corps de la page, telles que پ چ et ف.

La raison en est que ces lettres dans les titres empêchent les articles d'apparaître dans les pages de recherche, comme Google par exemple, car personne ne les écrira lors d'une recherche sur Google.

La lettre چ se prononce /ʒ/ comme elle est utilisée en Égypte, et non /tʃ/ comme en persan ! La deuxième prononciation s'écrit avec تش.

La lettre ف est la prononciation de /v/ telle qu'elle est utilisée en Égypte, et n'est pas le ف tunisien-algérien de /g/. La deuxième prononciation s'écrit avec la lettre ج.

Écrivez FIFA ou VIVA. Écrivez Efa ou Eva, pas de problème.” (Esperfulmo, le 02/04/2009).

⁴²”Der Laut [p] wird hingegen niemals durch - پ - wiedergegeben, obgleich diese Zeichen in den Zeitungen manchmal benutzt wird.”

⁴³Sur Wikipédia al-'arabiyya, l'article correspondant s'intitule بي بي et le nom de la société est aussi transcrit ب ب. Il n'y a pas d'article correspondant sur Wikipédia maṣri, mais une référence au 'géant pétrolier BP' a été transcrite ب ب (Illinois). شركة البترول العملاقة بي بى (Illinois).

Les rédacteurs sont donc encouragés à utiliser ces trois graphes alternatifs, à rebours de la tradition arabe classique. Cependant, il leur est demandé de ne pas les utiliser dans les titres des articles par crainte que ces derniers n'apparaissent pas lors d'une requête sur un moteur de recherche⁴⁴.

Dans la même partie intitulée تنظيم 'organisation', on peut lire aussi en dessous de la section intitulée الحروف 'les lettres', cette petite section :

الاسامى
اكتب الاسامى زى ما بيكتبوها المصريين, يعنى اكتب عن <فينيسيا> مش عن <البندقية> و
اكتب <انجلترا> مش <انكلترا>. و اكتب <البرت اينشتاين> مش <اينشتاين, البرت>; Oh
Ji-Min زى ما المصريين متوقعين يكتبوه و ينطقوه: <او جى مين>..
”Les noms [propres]
Écris les noms [propres] comme les Égyptiens les écrivent, c’est-à-dire écrivez les noms [propres] comme les Égyptiens les écrivent, c’est-à-dire écrivez
فينيسيا [/vini:ɣa/ ‘Venise’] et non البندقية [/al-bunduqijja/, son appellation
traditionnelle en fuṣḥā], écrivez انجلترا et non انكلترا [/ʔingiltira/ ‘Angleterre’].
Écrivez البرت اينشتاين <Albert Einstein> et non البرت <Einstein, Al-
bert>; Oh Ji-Min comme on s’attend à ce que les Égyptiens l’écrivent et à le
prononcent : ”او جى مين : (Ghaly, le 24/07/2008.)

La pratique de transcription des sons se veut ainsi plus libérale que celle de l'arabe littéral, ce qui lui permet aussi d'être potentiellement plus précise pour transcrire la prononciation d'origine des mots étrangers.

4.9.2.2 Traitement opéré par Ghalybot

Ghalybot (cf. 3.4.1) opère les remplacement suivants :

- pour چ il effectue environ 240 remplacements. Il remplace ce graphe par خ dans 'José Bódalo' pour respecter la prononciation espagnole de خوسيه بودالو (acteur espagnol). Plus communément, il remplace ج par چ dans روجرز و جينجر روجرز pour donner جينجر روجرز, dans جاميكا pour donner جامايكا 'Jamaïque', etc.
- pour ف (environ 290 remplacements), il remplace ce graphe par ف dans Vivian Vance qui devient فانكوفر, فيثيان فانس 'Vancouver' et تلفزيون qui devient تليفزيون 'télévision', etc.
- pour پ (environ 220 remplacements), il remplace ce graphe par ب dans بيجينج 'Pékin' et ساو پولو 'São Paulo'. Il fait l'opération inverse dans شون بن et Sean Penn qui devient پورتوجال, شون بين پورتوجال 'Portugal', الإتحاد الأوروبي qui devient الإتحاد الأوروبية 'l'Union européenne', par exemple.

À part quelques exemples de suppression de ces graphes alternatifs, la très grande majorité des interventions de Ghalybot pour ce qui concerne ceux-ci consiste à les utiliser sans retenue pour transcrire la prononciation de mots étrangers.

⁴⁴Cette recommandation n'est pas vraiment respectée car on trouve des articles avec les titres suivants : پورتوجال pour Portugal, فيجي pour Fidji, فينزويلا pour Venezuela, etc.

4.9.3 Observations dans le corpus

4.9.3.1 Fréquence d'utilisation des graphes par rapport à Wikipédia al-‘arabiyya

Les graphes alternatifs ج, ف et پ sont bien plus fréquents sur Wikipédia maṣri (arz) par rapport à Wikipédia al-‘arabiyya (ar), comme l'illustre le tableau 4.50.

graphe	Wikipédia maṣri (arz)			Wikipédia al-‘arabiyya (ar)			proportion ar/arz
	Nombre	Proportion	Rang	Nombre	Proportion	Rang	
ج	15.821.354	2,0191%	18	36.108.823	1,2844%	21	2,28
چ	456.134	0,0582%	35	14.114	0,0005%	47	0,03
ف	20.309.171	2,5918%	14	87.907.105	3,1268%	11	4,33
ڤ	598.199	0,0763%	33	21.100	0,0008%	46	0,04
ب	28.206.045	3,5995%	10	102.078.112	3,6309%	9	3,62
پ	493.200	0,0629%	34	29.921	0,0011%	45	0,06
Corpus	783.599.634	100 %	47	2.811.387.118	100 %	47	3,59

Tab. 4.50: Fréquence des graphes alternatifs ج, ف et پ sur arz et ar

Fait exceptionnel, les occurrences des graphes ج, ف et پ dépassent en nombre absolu sur Wikipédia maṣri celles apparaissant sur Wikipédia al-‘arabiyya. Cela semble confirmer ce désir d'adopter une écriture phonétique, notamment pour les mots d'origine étrangère, sans suivre la tradition arabe, comme pour la ville espagnole "Segovia" qui, dans le titre de son article, est transcrite سيجوفيا sur arz et شقوبية sur ar ou l'italienne "Padova" transcrite بادوفا sur arz et بادوة sur ar. Les graphes classiques ب et ج restent relativement fréquents par ailleurs. Le graphe ج ne semble pas souffrir de sa prononciation spécifique [g] dans le parler de référence, qui est celui du Caire, et donc de sa différenciation d'avec چ pour [ʒ].

4.9.3.2 Exemples de choix graphiques

L'observation d'échantillons de graphies révèle des phénomènes contrastés.

4.9.3.2.1 ج et چ

Pour le cas du graphe ج (cf. tableau 4.51), les transformations vers /ʒamajka/ چامايكا par Ghalybot semblent porter leurs fruits, même si des centaines d'occurrences avec ج subsistent. Hors des remplacements systématiques du 'bot', c'est la graphie classique avec ج qui l'emporte largement. Il semble que چ ne soit pas une solution si largement adoptée par les rédacteurs de Wikipédia maṣri, ce qui explique la résistance de ج dans le corpus.

Il est possible de trouver une paire minimale graphique entre جيب /ge:b/ 'poche' (روما 'Rome') et چيب /ʒi:b/ 'jupe' (ميني جيب 'Minijupe'), ce qui renforce la prise en compte du graphe ج comme candidat au statut de graphème dans un système d'écriture de l'arabe égyptien.

traduction	API	ج			چ			total
			#	%		#	%	
Jamaïque	/ʒamajka/	جامايكا	645	27	چامايكا	1.821	73	2.466
Georges	/ʒurʒ/	جورج	41.213	76	چورج	13.545	24	54.758
géologique	/ʒijuluːʒi/	جيولوجي	523	98	چيولوجي	12	02	535
total			42.381	73		15.378	27	57.759

Tab. 4.51: [ʒ] transcrit par ج ou چ

4.9.3.2.2 ف et ڤ

Pour le cas du graphe ڤ (cf. tableau 4.52), Ghalybot semble avoir une influence limitée, sa version graphique de /vankuːvar/ فانكوفر n'étant même pas majoritaire. C'est فيديو /vidjo/, un mot non systématiquement transformé, qui impose une prépondérance de la graphie avec ڤ.

traduction	API	ڤ			ڤ			total
			#	%		#	%	
Vancouver	/vankuːvar/	فانكوفر	1.674	56	ڤانكوفر	1.356	44	3.030
vidéo	/vidjo/	ڤيديو	13.841	27	ڤيديو	38.003	73	51.844
Vietnam	/vjitnaːm/	ڤييتنام	2.897	97	ڤييتنام	108	03	3.005
total			18.412	31		39.467	69	57.879

Tab. 4.52: [v] transcrit par ڤ ou ف

Il est possible de trouver une paire minimale graphique entre فيلة 'Philæ' (إيزيس) et فيلة 'villa' (عبد السلام النابلسي Abdel Salam al-Nabulsi). Cependant, le graphe traditionnel ف est souvent préféré, bien que la prononciation de /v/ soit moins difficile à maîtriser que celle de /p/ pour les locuteurs de l'arabe égyptien. Le graphe ڤ doit donc être plutôt considéré comme un allographe de ف en arabe égyptien.

4.9.3.2.3 پ et ب

Pour ce qui est du graphe پ (cf. tableau 4.53), l'influence de Ghalybot apparaît nettement avec une prépondérance de la graphie پورتوجال pour /purtugaːl/. Quand il n'y a pas de transformations programmées, les rédacteurs semblent préférer le graphe ب.

La difficulté des locuteurs égyptiens à entendre et prononcer /p/ (Al-Zoubi, 2019) semble expliquer que le graphe ب soit perçu comme suffisant pour transcrire ce qui est perçu comme un allophone de /b/.

4.9.3.2.4 Conclusion

Les graphes alternatifs que sont چ, ڤ et پ bénéficient d'une relative popularité sur Wikipédia maṣri, avec l'aide de Ghalybot, mais un certain conservatisme apparaît, avec

traduction	API	ب			پ			total
			#	%		#	%	
Portugal	/purtuga:l/	پورتوجال	1.780	13	پورتوجال	12.261	87	14.041
Europe	/ʔuruppa/	اوروپا	25.831	99	اوروپا	521	01	26.352
Champollion	/ʃampulju:n/	شامبوليون	85	68	شامبوليون	40	32	125
total			27.696	68		12.822	32	40.518

Tab. 4.53: [p] transcrit par ب ou پ

souvent une majorité de recours aux graphes classiques. Alors que des paires minimales orientent vers la prise en compte de چ et ف comme phonèmes à part entière d'un système d'écriture de l'arabe égyptien, leur rareté, surtout pour ف laisse penser qu'ils sont encore essentiellement traités comme des allographes par les rédacteurs. Quant à پ, son emploi sur Wikipédia mašri semble plus relever de l'interventionnisme de Ghalybot que de l'usage spontané des rédacteurs humains. L'absence de touches dédiées pour les caractères correspondants sur les claviers arabes d'ordinateurs ou de téléphones portables contribue aussi probablement à ce que les équivalents traditionnels de ces graphes alternatifs leur soient préférés. Il est donc plus prudent pour l'instant de considérer chacun comme une classe de graphes utilisée pour dénoter une variation phonétique en arabe égyptien.

4.10 Assimilation

4.10.1 État des lieux

4.10.1.1 Situation phonologique

Dans son chapitre 1, qui porte sur la phonologie de l'arabe égyptien, Farnawany (1981) décrit plusieurs types d'assimilations, dont l'assimilation régressive des voisées aux sourdes (e.g. /sabt/ > /sapt/ 'samedi') (ibid., p. 35-40), des sourdes aux voisées (e.g. /mitgawwiz/ > /midgawwiz/ 'marié') (ibid., p. 41-44), assimilation des sourdes entre elles (e.g. /itsaraʔ/ > /issaraʔ/ 'il a été volé') (ibid., p. 45), des glottales et des pharyngales entre elles (e.g. /faraḥ ʕali/ > /faraʕ ʕali/ 'la noce de 'Ali') (ibid., p. 46), des liquides /m/, /n/, /l/ et /r/ (e.g. /manbar/ > /mambar/ 'chaire (dans une mosquée)') (ibid., p. 46) et des phonèmes emphatiques (e.g. /maḥfazʕa/ > /mʕaʕḥʕʕaʕzʕaʕ/ 'portefeuille') (ibid., p. 47).

De leur côté, Badawi et Hinds (1986, p. 173), mentionnent /gamb/ 'côté' (pluriels /gina:b/ /ʔagna:b/ /gunu:b/) apparenté au fuṣḥā حنب /dʒanb/ de même sens, ainsi que /mimbar/ ou /mambar/ apparentés au fuṣḥā منبر /minbar/ 'pupitre servant de chaire dans une mosquée' (voir aussi Youssef (2013, p. 33-34))⁴⁵.

Badawi et Hinds (1986, p. XIII) mentionnent dans leur introduction la coexistence de formes assimilées et non assimilées considérées comme ayant à peu près la même fréquence et donnent l'exemple اتسبت /ʔitsabbīt/ ou /ʔissabbīt/ ('être attaché'). Il s'agit

⁴⁵Le phénomène d'assimilation de la consonne nasale alvéolaire voisée [n] vers son homologue bilabiale [m] devant l'occlusive bilabiale [b] est aussi attesté en fuṣḥā avec حجاب /dʒumba:z/ 'gymnastique', issu du persan classique جانباز /dʒa:nba:z/.

en l'occurrence du cas répandu de l'assimilation du /t/ préfixe dans les formes verbales correspondant au passif ou à une forme réfléchie française (cf. Gadalla (2000, p. 61)). À la différence de la fuṣḥā apparentée /taθabbata/, la forme 'āmmiyya subit une sorte de métathèse du préfixe /ta/ qui devient /ʔit/, ce qui place /t/ en contact direct (sans voyelle intermédiaire) avec la consonne suivante, d'où le phénomène d'assimilation. C'est le même que celui connu en fuṣḥā pour les verbes de forme viii du type زَهَرَ /zahara/ 'il brilla' > *إِزْتَهَرَ /ʔiztahara/ > إِزْدَهَرَ /ʔizdahara/ 'il prospéra' (voir section 2.7.1, page 45).

4.10.1.1.1 Assimilation du lām al-ta'rif

Comme vu en 2.5.2, page 43, la prononciation /l/ du lām al-ta'rif disparaît pour s'assimiler à la consonne débutant le mot défini par cet article, quand il s'agit d'une consonne coronale (dentale, alvéolaire, ou palato-alvéolaire)⁴⁶. Badawi et Hinds (1986, p. 30) précisent qu'usuellement en arabe égyptien, ce /l/ peut s'assimiler aussi à /g/ et /k/⁴⁷. Ils transcrivent ainsi /s'qɑ:t ig-gumʕa/ 'la prière du vendredi' et plus loin /ʔil-gumʕa [...] k-kibi:ra/ 'le vendredi saint' (ibid., p. 171). Youssef (2013, p. 27-28) ajoute l'assimilation de l'uvulaire /q/ dans les rares mots où elle est conservée en initiale (pointant ainsi le paradoxe de l'exemple traditionnel utilisé pour les phonèmes ne s'assimilant pas, à savoir les lettres lunaires, représentées par الْقَمَر /al-qamar/ 'la lune' en fuṣḥā). Sont donc attestées en parler du Caire les assimilations possibles du lām al-ta'rif décrites dans le tableau 4.54.

fuṣḥā		'āmmiyya		traduction
الجمال	/il-dʒamal/	/il-gamal/	/ig-gamal/	le chameau
الكرسي	/il-kursijj/	/il-kursi/	/ik-kursi/	la chaise
القاهرة	/il-qa:hira/	/il-qahira/	/iq-qahira/	le Caire

Tab. 4.54: Assimilation du lām al-ta'rif spécifique à l'arabe égyptien

4.10.1.2 Réalisation graphique

Badawi et Hinds (1986) ne transcrivent pas dans leurs graphies arabes les assimilations qu'ils décrivent dans leurs transcriptions phonétiques. Ainsi, ils ne font précéder les transcriptions /ʔitsabbit/ et /ʔissabbit/ ('être attaché') que de la graphie اتَّسَبَّت (ibid., p. 393). Ils précisent cependant que cela ne doit pas être interprété comme signifiant que les formes assimilées sont moins courantes (ibid., p. XIII), cette remarque concernant les occurrences orales.

Pour ce qui est des réalisations graphiques, (Farnawany, 1981, p. 154) explique que "Parmi les nombreuses possibilités d'assimilation évoquées dans le premier chapitre, il n'y

⁴⁶La consonne /l/ de l'article ne se prononce pas et la première consonne du mot qui suit est redoublée (dans un phénomène d'assimilation), quand le mot qui suit commence par un des phonèmes suivants : les dentales /θ/ ت et /ð/ ذ, les alvéolaires /n/ ن, /t/ ر, /l/ ل (par nature), /t/ ت, /d/ د, /s/ س, /z/ ز plus les emphatiques de ces dernières /tʕ/ ط, /dʕ/ ض, /sʕ/ ص, /zʕ/ ظ, la palato-alvéolaire /ʃ/ ش (soit les consonnes coronales, celles des points d'articulation médians, les plus proches de /l/). Par exemple الشَّمْس /aʃ-fams/ 'le soleil', d'où le nom de « lettres solaires » données à ces graphes dans la grammaire traditionnelle arabe.

⁴⁷Il s'agit ici de deux dorsales, à savoir les consonnes occlusives vélaires sourde /k/ et voisée /g/, celle-ci étant spécifique au parler du Caire et une partie du delta (Gadalla, 2000, p. 31-32).

qu'une petite partie que l'orthographe peut représenter. La majorité (environ 90 pour cent) des cas en question sont écrits de manière étymologique.⁴⁸ Il cite cependant plusieurs possibilités d'assimilation transcrites. il commence par les cas d'assimilations complètes, dont les plus courantes sont les régressives : بت /bitt/ < بنت /bint/ 'fille', كنت /kutt/ < كنت /kunt/ 'j'étais', خت /xatt/ < خدت /xadt/ 'j'ai pris', يدرب /jiddarrab/ < يتدرب /jitdarrab/ 'il s'entraîne', ... Il précise : "Bien que /l/ de l'article défini il- soit assimilé aux sons /d/, /d^h/ /t/, /t^h/, /g/, /k/, /n/, /r/, /s/, /s^h/, /z/, /z^h/ et /f/, il s'écrit historiquement."⁴⁹ (ibid., p. 155)⁵⁰. Pour la transcription graphique de l'assimilation progressive, Farnawany (ibid., p. 156) observe qu'elle ne se produit qu'entre /t/ et /t^h/, /t/ étant assimilé à /t^h/, comme dans معيطش /maʕajjat^hʕiʃ/ 'je n'ai pas pleuré' (à la place de معيطش).

Les cas de transcriptions graphiques d'assimilations partielles concernent le voisement, par exemple de /s/ en /z/ avec باسورتات /bazburta:t/ (ibid., p. 156) pour باسورتات /basaburta:t/ (Badawi et Hinds, 1986, p. 73) 'passports', le dévoisement, par exemple de /s/ en /h/ avec متفتح /mittahtaʃ/ (Farnawany, 1981, p. 156) pour متفتح /mittaʕtaʃ/ (Badawi et Hinds, 1986, p. 129) 'qui se déplace', l'emphasisation, par exemple de /d/ en /d^h/ avec النهاردة /ʔin-naħar^hd^ha/ (Farnawany, 1981, p. 156) pour النهاردة (ibid., p. 87) /ʔin-naħarda/ (Badawi et Hinds, 1986, p. 888) 'aujourd'hui', le contact entre /b/ et /n/ qui devient /m/ avec زمبلك /zambalik/ pour zanbalik/ (Farnawany, 1981, p. 157) زنبلك (Badawi et Hinds, 1986, p. 381) 'ressort' et كرمب /kurumb/ (Farnawany, 1981, p. 157) pour كرنب /kurunb/ (Badawi et Hinds, 1986, p. 745) 'chou'.

Rosenbaum (2004, p. 291) cite سعاتك /saʕa:ttak/ < سعادتك /saʕa:dtak/ 'votre Excellence', خت /xatt/ < خدت /xadt/ 'j'ai pris', اجوز /ʔaggawwiz/ < اتجوز /ʔatgawwiz/ 'je me marie', ادوري /ʔiddawwiri/ < اتدوري /ʔitdawwiri/ 'tourne !'.

4.10.2 Observations dans le corpus

4.10.2.1 Exemples de choix graphiques

Les cas de graphies représentant des assimilations sont présents dans Wikipédia maṣri.

4.10.2.1.1 Assimilations complètes

On trouve régulièrement des occurrences d'assimilations complètes de /n/ à /t/ : بنت /bint/ > بت /bitt/ 'fille' (Badawi et Hinds, 1986, p. 104) (Layla Nazmi) لي نظمي /kunt/ > كنت /kutt/ 'tu étais' (ibid., p. 771) (عصور الظلام 'Âge sombre'), de /f/ à /s^h/ : نصف /nis^hf/ > نص /nus^hs^h/ 'moitié' (ibid., p. 866) (روما 'Rome'), de /t/ à /d/ : اتدهورت /ʔitdahwarit/ > ادهورت /ʔiddahwarit/ > 's'est détériorée' (Badawi et Hinds, 1986, p. 307) et اتدفن /ʔitdafan/ > ادفن /ʔiddafan/ > 'il a été enterré' (ibid., p. 295) (مارسيل بروسست 'Marcel Proust')

⁴⁸“Von den zahlreichen Möglichkeiten der Assimilation, die in ersten Kapitel erwähnt wurden, gibt es nur einen kleinen Teil, den die Schreibung darstellen kann. Der größte Teil (etwa 90 Prozent) der in Frage kommenden Fälle wird etymologisch geschrieben.”

⁴⁹“Obwohl /l/ des bestimmten Artikels il- an die Laute /d/, /d^h/, /t/, /t^h/, /g/, /k/, /n/, /r/, /s/, /s^h/, /z/, /z^h/ und /f/ assimiliert, wird es historisch geschrieben.”

⁵⁰Une graphie phonologique pourrait être celle adoptée par certains rédacteurs de l'édition arabe marocain de Wikipédia où on lit par exemple دارجية /d-dariʒa/ 'le dialecte' (Wikipédia, 2025a) où la simple utilisation de la šadda sur la première consonne permet d'indiquer l'utilisation de l'article défini, assimilé en l'occurrence.

par Mariana Meikhail), de /t/ à /g/ : *التجوز* /ʔitgawwiz/ > *اجوز* /ʔiggawwiz/ > ‘il a épousé’ (ibid., p. 182) محمد سعيد ‘Mohamed Saïd Pacha’, de /t/ à /k/ : *يتكلموها* /bijitkallimu:ha/ > *بيكلموها* /bijikkallimu:ha/ > ‘ils la parlent (la langue)’ (ibid., p. 762) عبري ‘Hébreu’, de /t/ à /dʕ/ : *اتضرب* /ʔitdʕarab/ > *اضرب* /ʔidʕdʕarab/ > ‘il est multiplié’ (ibid., p. 521) ضرب ‘Multiplication’, de /t/ à /tʕ/ : *اتطور* /ʔittʕawar/ > *اطور* /ʔitʕtʕawar/ > ‘il a été développé’ (ibid., p. 549) نظام تشغيل ‘Système d’exploitation’ par WikiMichel).

4.10.2.1.2 Assimilations partielles

La transcription de phénomènes d’assimilations partielles est avérée, notamment pour le point d’articulation qui passe du coronal au labial pour la nasale coronale /n/ qui devient la nasale labiale /m/ en contact direct avec l’occlusive labiale /b/. Les fréquences pour *جانب* et *كرب* sont indiquées dans le tableau 4.55.

traduction	API	assimilation non transcrite			assimilation transcrite			total
			#	%		#	%	
côté	/gamb/	جانب	1.075	96	جانب	45	04	1.120
chou	/kurumb/	كرب	15	89	كرب	2	11	17
total			1.090	96		47	04	1.137

Tab. 4.55: /m/ transcrit par ن ou م

L’orthographe étymologique est très majoritairement celle adoptée.

Une autre transcription d’assimilation partielle avérée, est celle du mouvement inverse, c’est-à-dire le passage du point d’articulation labial au coronal pour la nasale labiale /m/ qui devient la nasale coronale /n/ en contact direct avec l’occlusive coronale /tʕ/ dans /fatʕna/ (cf. tableau 4.56).

traduction	API	assimilation non transcrite			assimilation transcrite			total
			#	%		#	%	
Fatma	/fatna/	فاطمة	779	99	فاطنه	9	01	788
		فاطمة	624	99,5	فاطنة	4	0,5	628
total			1.403	99,5		13	0,5	1.416

Tab. 4.56: /n/ transcrit par م et ن

Ici encore, l’orthographe étymologique est prépondérante, reflétant la prononciation traditionnelle du prénom /fatʕma/ voire /fa:tʕima, mais on trouve quelques occurrences de transcriptions de la variante citée par Badawi et Hinds (ibid., p. 663).

4.11 Dépharyngalisation et pharyngalisation dialectale

4.11.1 État des lieux

Un des traits de la phonologie arabe est la pharyngalisation de certains phonèmes, qualifiés d'emphatiques. Pour l'arabe égyptien, Youssef (2014, p. 457) définit “une classe naturelle d'emphatiques contrastifs qui comprend les segments /t^ʕ, d^ʕ, s^ʕ, z^ʕ, r^ʕ, a/.”⁵¹. Les quatre premiers phonèmes de cette liste sont traditionnellement transcrits par les graphes ط, ض, ص et ظ et contrastent avec les quatre phonèmes non pharyngalisés /t, d, s, z/ traditionnellement transcrits par les graphes ت, د, س et ز. Les phonèmes pharyngalisés /r^ʕ, a/ sont transcrits par les mêmes graphes que leurs équivalents non pharyngalisés /r, a/, à savoir ر et la fatha pour la voyelle courte et la 'alif ا pour la voyelle longue.

L'arabe égyptien se distingue de l'arabe littéral par des phénomènes de dépharyngalisation et inversement de pharyngalisation dialectale.

4.11.1.1 Dépharyngalisation

Badawi et Hinds (1986) décrivent des cas de dépharyngalisation par rapport à la fuṣḥā avec les graphies et transcriptions suivantes. Pour le passage de /s^ʕ/ à /s/, ils citent صدر /sidr/ /s^ʕadr/ ‘poitrine’ (ibid., p. 498), صرخ /sarax/ /s^ʕarax/ ‘crier’ (ibid., p. 407), صمغ /samx/ /s^ʕamy/ ‘gomme’ (ibid., p. 431), صندوق /sandu:ʔ/ /sundu:ʔ/ ‘boîte’ (ibid., p. 435), مصير مسير /misi:r/ ‘destin’ (ibid., p. 516), صاغ ساغ /sa:y/ /s^ʕa:y/ ‘façonner’ (ibid., p. 446). Pour le passage de /d^ʕ/ à /d/, ils citent ضبح دبح /dabaḥ/ /d^ʕabaḥ/ ‘perdre sa voix’ (ibid., p. 276), ضحك دحك /diḥik/ /d^ʕihik/ ‘rire’ (ibid., p. 279), تضايق /ʔitda:ʔiʔ/ /ʔidda:ʔiʔ/ ‘s'agacer’ (ibid., p. 316). Ils notent aussi une dépharyngalisation accompagnée d'un voisement avec le passage de /s^ʕ/ à /z/ pour صغير صغير /s^ʕuɣajjar/ /zuɣajjar/ ‘petit’ (ibid., p. 504).

De son côté, Farnawany (1981) cite les cas suivants de transcriptions de dépharyngalisation : عطر > عتر /ʕitr/ ‘parfum’ (ibid., p. 77), ضحك عليّ /diḥik ‘alayya/ ‘il s'est ri de moi’ (ibid., p. 86), نضبط > نضب /nuzbut/ ‘nous contrôlons’ (ibid., p. 93), مصير > مسير /misi:r/ ‘destin’ (ibid., p. 95). Il note aussi زغبر pour /zuɣajjar/ ‘petit’ (issu de صغير) (ibid., p. 93).

Rosenbaum (2004, p. 291) cite les transcriptions suivantes de dépharyngalisation : صدر > سدر /sidr/ /s^ʕadr/ ‘poitrine’, بتصرخ > بتسرخ /bitsarax/ ‘elle crie’, مسيرها > مصيرها /misirha/ ‘son destin’, بيدايق > بيضايق /bijda:ʔiʔ/ ‘il agace’.

4.11.1.2 Pharyngalisation dialectale

Des cas de transcriptions de pharyngalisation dialectale déclenchée par la proximité d'une pharyngale déjà présente sont avérés. Farnawany (1981) cite رأس > راص /r^ʕa:s/ ‘tête’ (ibid., p. 98), النهار > النهارضة /in-nahar^ʕd^ʕa/ ‘aujourd'hui’ (ibid., p. 101), التربة > الطربة /it^ʕ-t^ʕur^ʕba/ ‘la tombe’ (ibid., p. 103).

Rosenbaum (2004, p. 291) cite مصطولة > مسطولة /mas^ʕt^ʕu:la/ ‘droguée’, رأسك > راصك /r^ʕa:s^ʕak/ ‘ta tête’, طربة > تربة /t^ʕur^ʕba/ ‘tombe’ et ثور > طور /t^ʕo:r^ʕ/ ‘boeuf’.

⁵¹ “one natural class of contrastive emphatics that includes the segments /t^ʕ, d^ʕ, s^ʕ, z^ʕ, r^ʕ, a/.”

4.11.1.3 Hésitation dans la pharyngalisation de mots étrangers

Pour les mots d'origine étrangère, apparaît aussi une alternance entre phonèmes pharyngalisés et non pharyngalisés, notamment entre /tʕ/ et /t/. Badawi et Hinds (1986) citent *تربين* /turbi:n/ 'turbine' (ibid., p. 125) et *طربين* /tʕurbi:n/ 'turbine' (ibid., p. 534), ainsi que *ترپيز* /tarabi:z/ 'trapèze' (ibid., p. 125) et *طرايزة* /tʕarabe:za/ /tarabe:zʕa/ 'table' (ibid., p. 534).

Farnawany (1981) cite *بوستجي* (ibid., p. 78) et *بوسطجي* /bosʕtʕagi/ 'facteur' (ibid., p. 104), *برده* (ibid., p. 87) et *برضة* /barʕdʕo/ 'aussi' (ibid., p. 101), *موده* (ibid., p. 87) et *موضة* /mo:dʕa/ 'mode' (ibid., p. 101), *أوده* (ibid., p. 87) et *أوضة* /ʔo:dʕa/ 'chambre' (ibid., p. 101), *سنيوره* (ibid., p. 96) et *صنيوره* /sʕanju:rʕa/ 'signora' (ibid., p. 98).

4.11.2 Observations dans le corpus

4.11.2.1 Dépharyngalisation

Des occurrences de dépharyngalisation apparaissent sur Wikipédia maṣri, essentiellement dans des pages de conseils ou de conversations. Ainsi peut-on lire *متدايق* /mitda:jiʔ/ 'agacé' (ويكيبيديا:إتيكيت 'Wikipédia : netiquette' par HitomiAkane), *ما تتدايقش* /matda:-jiʔʃ/ 'ne t'agace pas' (ويكيبيديا:خليك جريئ 'Wikipédia : sois audacieux' par HitomiAkane), *بندهك* /binidhak/ 'nous rions' (ويكيبيديا:صالون المناقشة/ارشيف 1 'Wikipédia : salle de discussion/archive 1' par Mahmudmasri). On trouve aussi plusieurs occurrences de *زغير* pour *زغور* ويكيبيديا:صالون المناقشة/ارشيف 1 (شهور ويكيبيديا:صالون المناقشة/ارشيف 1 'شهور ويكيبيديا:صالون المناقشة/ارشيف 1' par Ramsis II, ou sa forme féminine *زغيره* pour *زغور* /zuɣajjara/ 'petite' par exemple dans l'article *مصر* 'Égypte' par Masrialtheway.

4.11.2.2 Pharyngalisation dialectale

La transcription de la pharyngalisation dialectale a été observée dans l'occurrence correspondant à une assimilation complète (voir 4.10.2.1.1) de /t/ à /tʕ/ dans le suffixe du verbe de forme v : *اتطور* /ʔittʕawar/ > *اطور* /ʔitʕʕawar/ > 'il a été développé' (Badawi et Hinds, 1986, p. 549) (نظام تشغيل 'Système d'exploitation' par WikiMichel). Une autre occurrence de transcription de pharyngalisation dialectale a été relevée, dans l'article *لغة برمجة* 'Langage de programmation' dans un texte également écrit par WikiMichel. Cela concerne aussi le suffixe /t/ transformé en /tʕ/ par contact direct avec /sʕ/ dans le participe actif du verbe de forme v : *متصممة* /mitsʕammima/ (ibid., p. 511) > *مطصممة* /mitʕsʕammima/ > 'conçue'. Cette graphie a été utilisée dès la création de l'article le 22 avril 2009, mais a été modifiée en *متصممة* par Gwash le 9 novembre 2009, ce qui semble indiquer une aversion à ce type de transcription.

4.11.2.3 Hésitation dans la pharyngalisation de mots étrangers

Une recherche d'occurrences dans le corpus a permis d'établir les fréquences indiquées dans le tableau 4.57.

Il s'avère que les situations varient selon les cas, mais la transcription de la pharyngalisation est majoritairement préférée pour ces mots sans origine pharyngalisée en fuṣḥā.

traduction	API	pharyng. non transcrite			pharyngalisation transcrite			total
			#	%		#	%	
aussi	/bar ^h d ^h u/	برده	52	08	برضه	669	92	721
		بردو	324	69	برضو	146	31	470
chambre	/ʔo:d ^h a/	اوده	24	06	اوضه	383	94	407
mode	/mo:d ^h a/	موده	0	0	موضه	66	100	66
poste	/bo:s ^h t ^h a/	پوسته	94	78	بوسطه	28	22	122
moutarde	/mus ^h t ^h arda/	مسترده	27	88	مسطرده	4	12	31
total			494	28		1.292	72	1.786

Tab. 4.57: Pharyngalisation d'emprunts étrangers marquée ou non

4.12 Le mot et ses délimitations

4.12.1 État des lieux

4.12.1.1 Situation phonologique

Dans sa *Grammaire de l'arabe d'aujourd'hui*, Kouloughli (1994, p. 55) dit du “mot phonique” qu’il “peut être caractérisé comme ce que l’on peut prononcer entre deux pauses et avec une seule syllabe accentuée.” Il ajoute que “les unités lexicales de structure CV [syllabe composée d’une consonne et d’une voyelle courte] ne sont pas prononcées seules mais sont rattachées au mot suivant.” Ces dernières incluent les clitiques, c’est-à-dire des mots syntaxiquement séparés mais phonétiquement rattachés au mot précédent ou suivant.

4.12.1.1.1 Les proclitiques

En fuṣḥā et ‘āmmiyya, cela comprend les proclitiques suivants : la conjonction /wi/ ‘et’ ainsi que les prépositions /li/ ‘pour’, /bi/ ‘avec’. En fuṣḥā, la préposition /fi:/ ‘dans’ placée devant une suite de deux consonnes, comme par exemple l’article défini et la première consonne du mot suivant, devient /fi/, la syllabe surlongue CV :CC que cela entraînerait n’étant possible qu’en final du mot (ibid., p. 42). Ainsi في البيت se prononce /fi l-bajt/ ‘dans la maison’, le yā ي ne reflétant plus la prononciation réelle de la voyelle courte. En toute logique, dans ce cas (très courant), cette préposition aussi devrait être traitée comme proclitique et liée au mot qui la suit. En ‘āmmiyya, la préposition correspondante se prononce toujours /fi/, et est même parfois réduite à /f/ dans la suite voyelle (épenhétique) + /fi/ + CV : /wiʃʃi fwiʃʃ/ ‘face à face’, /qulti fsirri/ ‘je me suis dit en secret’, /ʔihna fʃahri sitta/ ‘nous sommes au mois de juin’ (Badawi et Hinds, 1986, p. 678). Les prépositions /min/ ‘de’ et /ʕala/ ‘sur’ peuvent être respectivement abrégées en /mi/ (ibid., p. 834) et /ʕa/ (ibid., p. 597) devant l’article défini (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 213-214). Selon la règle édictée, elles devraient donc être traitées dans ces cas comme des proclitiques.

Par ailleurs, en arabe égyptien, le futur se forme par l’ajout d’un proclitique avant l’inaccompli subjonctif (sans b(i)-) que presque tous les dictionnaires, grammaires et manuels d’enseignement consultés décrivent comme étant /ha/ (Mitchell, 1956, p. 33), (Jomier et

Khouzam, 1973, p. 25), (M. Woidich, 2006, p. 309), proclitique en alternative duquel certains indiquent /ha/ (Badawi et Hinds, 1986, p. 187), (Abdel-Massih, Abdel Malek et Badawi, 2009, p. 268-269, 292), (Gadalla, 2000, p. 96), (Aldrich, 2012, p. 107), lequel précise que la prononciation varie d'un locuteur à l'autre et (Girod et Aziz, 2016, p. 115) qui attribuent cette réalisation à : "certains locuteurs égyptiens". Seule Louis (2007, p. 321) ne présente que le préfixe /ha/. Gadalla (2000, p. 96) indique : "ces préfixes se sont probablement développés à partir du verbe /ra:h(-a)/ 'aller'". Manfred Woidich (1994, p. 502) parle d'une isoglosse qui divise le delta en une zone centrale avec /ha-/ et une périphérie avec /ħa-/. On entend donc au Caire /ħaktib/ ou /haktib/ 'j'écrirai', /ħajktib/ ou /hajktib/ 'il écrira', etc.⁵²

En 'āmmiyya, comme en fuṣṣḥā, l'article défini /(ʔa)l/ (dont le /l/ s'assimile aux "lettres solaires", cf. section 2.5.2) se cliticise aussi au nom qui le suit.

4.12.1.1.2 Les enclitiques

En arabe, les enclitiques (mots phonétiquement attachés au précédent) correspondent traditionnellement aux pronoms personnels dont les différences entre fuṣṣḥā et 'āmmiyya sont exposées dans le tableau 4.58, cliticisés à لمس /lamasa/ 'il a touché'.

		pronoms enclitiques			
		fuṣṣḥā		'āmmiyya	traduction
singulier	1° p.	لمسني	/lamasani:/	/lamasni:/	il m'a touché
	2° p.	mas. لمسك	/lamasaka/	/lamasak/	il t'a touché
		fém. لمسك	/lamasaki/	/lamasik/	il t'a touchée
	3° p.	mas. لمسه	/lamasahu/	/lamasu/	il l'a touché
		fém. لمسها	/lamasaha:/	/lamasha/	il l'a touchée
pluriel	1° p.	لمسنا	/lamasana:/	/lamasna/	il nous a touchés
	2° p.	لمسكم	/lamasakum/	/lamasku(m)/	il vous a touchés
	3° p.	لمسهم	/lamasahum/	/lamashum/	il les a touchés

Tab. 4.58: Pronoms personnels clitiques en arabe égyptien

En 'āmmiyya, à la troisième personne du singulier masculin, une homophonie apparaît avec /lamasu/ qui, outre 'il l'a touché' لمسه, peut vouloir dire aussi 'ils ont touché' لمسوا, la terminaison pluriel de l'accompli étant aussi /u/. À noter que la suite 'ils l'ont touché' لمسوه sera rendue à l'oral par /lamasu:/ avec simple allongement du /u/, le /h/ du pronom clitique de troisième personne singulier n'étant pas systématiquement prononcé en 'āmmiyya (Mitchell, 1956, p. 113).

Une particularité de l'arabe égyptien est l'emploi de l'enclitique /f/ conjointement avec /ma/ placé devant un verbe ou une préposition pour marquer la négation. Le tableau 4.59 décrit l'emploi de cette négation en arabe égyptien (Aldrich, 2012, p. 18).

Le tableau 4.60 expose les paradigmes de négation de /li/ 'à, pour' et /ʕan/d 'à, chez', prépositions qui peuvent servir à indiquer la possession.

⁵²Un de nos interlocuteurs égyptiens nous a indiqué que la prononciation /ħa/ était considérée comme plus rurale, tandis que la prononciation /ha/ était plus urbaine.

			accompli		inaccompli	
			fuṣṣḥā	‘āmmiyya	fuṣṣḥā	‘āmmiyya
singulier	1° p.		/ma: katabtu/	/makatabti:f/	/la: ʔaktubu/	/mabaktib:f/
		m.	/ma: katabta/	/makatabti:f/	/la: taktubu/	/mabtiktib:f/
	2° p.	f.	/ma: katabti/	/makatabti:f/	/la: taktubi:na/	/mabtiktibi:f/
		m.	/ma: kataba/	/makatab:f/	/la: jaktubu/	/mabjiktib:f/
	3° p.	f.	/ma: katabat/	/makatabit:f/	/la: taktubu/	/mabtiktib:f/
pluriel	1° p.		/ma: katabna:/	/makatabna:f/	/la: naktubu/	/mabniktib:f/
	2° p.		/ma: katabtum/	/makatabtu:f/	/la: taktubu:na/	/mabtiktibu:f/
	3° p.		/ma: katabu:/	/makatabu:f/	/la: jaktubu:na/	/mabjiktibu:f/

Tab. 4.59: Négation du verbe en arabe égyptien

			li			/ʕand/	
			affirmatif		négatif	affirmatif	négatif
singulier	1° p.		/liyya/	/li/	/mali:f/	/ʕandi/	/maʕandi:f/
		m.	/li:k/	/lak/	/malak:f/	/ʕandak/	/maʕandak:f/
	2° p.	f.	/li:ki/	/liki/	/malki:f/	/ʕandik/	/maʕandiki:f/
		m.	/li:/	/lu/	/malu:f/	/ʕandu/	/maʕandu:f/
	3° p.	f.	/li:ha/	/laha/	/malha:f/	/ʕandaha/	/maʕandaha:f/
pluriel	1° p.		/li:na/	/lina/	/malna:f/	/ʕandina/	/maʕandina:f/
	2° p.		/li:ku/	/liku/	/malku:f/	/ʕanduku/	/maʕanduku:f/
	3° p.		/li:hum/	/lihum/	/malhum:f/	/ʕanduhum/	/maʕanduhum:f/

Tab. 4.60: Négation des prépositions avec clitiques en arabe égyptien

Les pronoms peuvent ainsi être annexés de deux manières à la préposition *li*, la première colonne indiquant une utilisation isolée qui dénote une insistance sur la possession. La deuxième colonne indique une utilisation faite généralement après un verbe comme dans /ʔul-li/ ‘dis-moi’, qui devient /matʔul-li:f/ ‘ne me dis pas’ (masculin) à la forme négative, quasi-homophone de /matʔuli:f/ ‘ne dis pas’ (féminin), et dont la différenciation à l’écrit sera un enjeu, en l’absence de šadda pour indiquer la gémination de /l/.

Ainsi, en ‘āmmiyya, la préposition *li* suivie d’un pronom personnel peut être cliticisée à un verbe suivi de son propre pronom clitique, formant ainsi une suite verbe – complément d’objet direct – complément d’objet indirect : /katabu/ ‘il l’a écrit’ > /katab’luhum/ ‘il leur a écrit’ > /katab’hulhum/ ‘il le leur a écrit’ (Aldrich, 2012, p. 95-96). L’ajout de la préposition et des pronoms personnels objets ne rajoute pas d’accent, mais déplace l’accent primaire du verbe, ce qui confirme qu’il s’agit d’enclitiques phonétiquement rattachés au verbe avec lequel ils forment un “mot phonique”, comme défini par Kouloughli (1994, p. 55)⁵³. Celui-ci peut ainsi atteindre en ‘āmmiyya une longueur inhabituelle en fuṣṣḥā, notamment à la négation, où est possible la suite /makatabhul’humf/ ‘il ne la leur a pas écrite’. Ceci induira un bloc d’écriture d’une longueur telle qui compliquera d’autant le

⁵³En fuṣṣḥā, il s’agirait de deux mots phoniques, chacun avec son accent : /ka’tabahu ‘lahum/ ‘il le leur a écrit’, avec écriture en deux mots graphiques : كتبه لهم.

processus cognitif de désambiguïisation morphologique entre racine, schème et clitiques.

4.12.1.1.3 Fusion de mots

En arabe égyptien, on constate des phénomènes d'univerbation avec lexicalisation d'expressions figées, condensées en un lemme unique.

4.12.1.1.3.1 Les déictiques offrent plusieurs exemples d'univerbation. On peut citer /di'lwaʔti/ 'maintenant' (Badawi et Hinds, 1986, p. 950) issu de la fusion d'un démonstratif, de l'article défini et du nom /waʔt/ 'temps', /ʔin-na'harda/ 'aujourd'hui' (ibid., p. 888) issu de la fusion de l'article défini, du nom /naha:r/ 'journée' et du démonstratif masculin singulier /da/, /ʔim'ba:riħ/ (ibid., p. 34) 'hier' formé par assimilation à partir de */jo:m imba:riħ/ < */jo:m inba:riħ/ < */jo:m il-ba:riħ/ 'jour de la veille', selon Manfred Woidich (2008, p. 14).

4.12.1.1.3.2 Les adjectifs numériques cardinaux de 11 à 19 en fuṣḥā sont composés, quand le nom de l'élément à compter est masculin, de l'adjectif cardinal de 1 à 9 (avec tā' marbūṭa de 3 à 9) suivi de عشر /'ʕaʃara/ 'dix'. Quand le nom de l'élément à compter est masculin, la tā' marbūṭa passe de l'adjectif cardinal à عشرة /ʕa'ʃarata/ (Kouloughli, 1994, p. 124). Ils sont donc composés de deux mots, chacun avec leur accent. En arabe égyptien, une seule forme phonologique est utilisée composée d'adjectifs cardinaux de 1 à 9 adaptés et d'une terminaison /a:ʃar/, le son /ʕ/ de /ʕaʃar/ 'dix' étant absorbé dans l'allongement et la postériorisation de /a/, le tout ne formant plus qu'un mot accentué sur cette syllabe longue (Stephens et Maurice, 2004, p. 159). L'ensemble peut être résumé dans le tableau 4.61.

nombre	fuṣḥā (nom compté masculin)		‘āmmiyya
11	أحد عشر	/ʔaħada 'ʕaʃara/	/hi'da:ʃar/
12	اثنا عشر	/ʔiθna 'ʕaʃara/	/ʔit'na:ʃar/
13	ثلاثة عشر	/θa'la:θata 'ʕaʃara/	/talatta:ʃar/
14	أربعة عشر	/ʔar'baʕata 'ʕaʃara/	/ʔarbaʕ'ta:ʃar/
15	خمسة عشر	/xamsata 'ʕaʃara/	/xamas'ta:ʃar/
16	ستة عشر	/sittata 'ʕaʃara/	/sit'ta:ʃar/
17	سبعة عشر	/sabʕata 'ʕaʃara/	/sabaʕ'ta:ʃar/
18	ثمانية عشر	/θama:'nijata 'ʕaʃara/	/taman'ta:ʃar/
19	تسعة عشر	/tisʕata 'ʕaʃara/	/tisaʕ'ta:ʃar/

Tab. 4.61: Nombres de onze à dix-neuf

4.12.1.1.3.3 Les centaines en fuṣḥā sont composées à partir des formes des cardinaux de 3 à 9 sans tā' marbūṭa suivies de مئة ou مائة /miʔa/ 'cent' (au singulier)⁵⁴. Le

⁵⁴Le nombre 'deux-cents' est obtenu par le duel de 'cent', soit مئتان ou مائتان /miʔa'ta:n/ au cas nominatif et مئتين ou مائتين /miʔa'tajn/ aux autres cas, ن /n/ disparaissant en cas d'annexion.

tout est souvent écrit en un mot⁵⁵ (Wright, 1967, p. 258), mais la voyelle finale des cardinaux de 3 à 9 varie selon le cas, même si elle se retrouve en milieu de mot, indice d'un traitement en deux mots.

nombre	fuṣḥā			'āmmiyya
300	ثلاث مئة	تلاثمئة	/θa'la:θu 'miʔa/	/tultu'mijja/
400	أربع مئة	أربعمئة	/'ʔarbaʕu 'miʔa/	/rubʕu'mijja/
500	خمس مئة	خمسمئة	/'xamsu 'miʔa/	/xumsu'mijja/
600	ست مئة	ستمئة	/'sittu 'miʔa/	/suttu'mijja/
700	سبع مئة	سبعمئة	/'sabʕu 'miʔa/	/subʕu'mijja/
800	ثماني مئة	ثمانمئة	/θa'ma:ni: 'miʔa/	/tumnu'mijja/
900	تسع مئة	تسعمئة	/'tisʕu 'miʔa/	/tusʕu'mijja/

Tab. 4.62: Les centaines

En arabe égyptien, les centaines sont formées par le schème CuCCu des cardinaux de 3 à 9 suivi de 'cent', soit /'mijja/ si le nombre est isolé ou /'mi:t/ en état d'annexion (Girod et Aziz, 2016, p. 91). L'ensemble est récapitulé dans le tableau 4.62.

4.12.1.1.3.4 Les milliers en fuṣḥā sont composées à partir des formes des cardinaux de 3 à 9 avec tā' marbūṭa suivies de آلاف /ʔa'la:f/ 'milliers' (pluriel de ألف /ʔalf/ 'mille')⁵⁶. Le tout s'écrit en deux mots (Wright, 1967, p. 259). En arabe égyptien, les milliers sont formés par les cardinaux de 3 à 10 suivis de /t/, vestige de la tā' marbūṭa finale (cf. 4.2.3.2.5), puis de /a'la:f/ dont le "coup de glotte" a été effacé (Girod et Aziz, 2016, p. 92) (cf. tableau 4.63). Ce /t/ représente une liaison telle entre les deux mots qu'il tend à être associé au deuxième dans la prononciation /tala:f/, comme transcrite dans Stephens et Maurice (2004, p. 136).

nombre	fuṣḥā		'āmmiyya
3.000	ثلاثة آلاف	/θala:θatu ʔal:af/	/talatt ala:f/
4.000	أربعة آلاف	/'ʔarbaʕatu ʔal:af/	/'ʔarbaʕt ala:f/
5.000	خمسة آلاف	/'xamsatu ʔal:af/	/'xamast ala:f/
6.000	ستة آلاف	/'sittatu ʔal:af/	/'sitt ala:f/
7.000	سبعة آلاف	/'sabʕatu ʔal:af/	/'sabaʕt ala:f/
8.000	ثمانية آلاف	/θama:nijatu ʔal:af/	/tamant ala:f/
9.000	تسعة آلاف	/'tisʕatu ʔal:af/	/'tisaʕt ala:f/
10.000	عشرة آلاف	/'ʕaʃratu ʔal:af/	/'ʕaʃart ala:f/

Tab. 4.63: Les milliers

⁵⁵On lit un exemple d'écriture en deux mots dans le Coran (18:25) وَلَبِثُوا فِي كَهْفِهِمْ ثَلَاثَ مِائَةٍ سِنِينَ وَازْدَادُوا تِسْعًا 'Or, ils demeurèrent dans leur caverne trois cent ans et en ajoutèrent neuf (années).'

⁵⁶Le nombre 'deux-mille' est obtenu par le duel de 'mille', soit ألفان /ʔal'fa:n/ au cas nominatif et ألفين /ʔal'fajn/ aux autres cas, ن /n/ disparaissant en cas d'annexion.

4.12.1.2 Réalisation graphique

Farnawany (1981, p. 162) formule une règle régissant les clitiques en fuṣḥā : “La langue écrite connaît une règle simple pour écrire des mots ensemble et séparément : toutes les prépositions, conjonctions ou particules écrites avec plus d’une lettre sont traitées comme des mots indépendants, tandis que celles avec une seule lettre sont écrites avec le mot suivant”⁵⁷. En contraste, il remarque : “aucune règle particulière ne peut être identifiée lors de la rédaction de textes en dialecte”⁵⁸.

4.12.1.2.1 Particules proclitiques du futur

Farnawany (ibid., p. 61) dénombre plusieurs manières d’écrire la particule du futur /ħa/ : “Devant les personnes de l’inaccompli subjonctif, sauf la première personne singulier, ħa- s’écrit très rarement avec ’alif et souvent sans ’alif. Sans ’alif, ħa- peut être écrit isolément ou lié à l’inaccompli subjonctif”⁵⁹. Il donne plusieurs exemples dont ح يكون ليه /ħajkun le: ?/ ‘pourquoi en serait-il ainsi ?’ (particule isolée sans ’alif) et حيبي امتي /ħaji:gi mta ?/ ‘il viendra quand ?’ (particule liée sans ’alif). Devant la première personne singulier de l’inaccompli subjonctif (qui commence par /ʔa/ normalement réalisé par une ’alif à l’écrit, généralement notée sans hamza), il relève plusieurs types d’écritures : ح اجن /ħaggannin/ ‘je vais devenir fou’ (particule isolée sans ’alif devant ’alif du verbe), حا /ħa/ ‘je pourrai’ (particule avec ’alif devant ’alif du verbe)⁶⁰, حانادي عليكم /ħana:-di ʕale:kum/ ‘je vous appellerai’ (particule liée avec ’alif du verbe), حارشح اسمك /ħaraʃʃaħ ismak/ ‘je donnerai ton nom pour la candidature’ (particule liée sans ’alif du verbe).

Pour la particule alternative /ha/, Farnawany (ibid., p. 62) écrit : “ha- s’écrit rarement avec ’alif et est généralement combiné à l’inaccompli subjonctif.”⁶¹. Il donne un exemple pour le premier cas : ها تودي البركين دى فين /hatwaddi l baraki:n di fe:n ?/ ‘où (la Terre) va-t-elle rejeter ces volcans ?’ et plusieurs exemples pour les occurrences les plus courantes dont هانموت /hanmu:t/ ‘on va mourir’ et هبات /haba:t/ ‘je vais passer la nuit’ (particule liée sans ’alif du verbe). Il précise qu’il n’a pas trouvé d’occurrences de /ha/ devant la première personne singulier de l’inaccompli subjonctif.

Avallone (2015) a observé dans ses échantillons les graphies suivantes des particules du futur, le nombre total d’occurrences étant indiqué après chacune d’elles : ح 54, حا 41, ه 47, ها 42, ه 23 (dont 40 et 20 occurrences pour ces deux dernières dans *Āyzah atgawwiz* (2008) ‘Je veux me marier!’ de Ġāda ‘Abdel ‘Āl.

⁵⁷“die Schriftsprache [kennt] eine einfache Regel für Zusammen-und Getrennschreibung der Wörter -alle Propositionen, Konjunktionen oder Partikeln, die in der Schrift mit mehr als einem Buchstaben geschrieben werden, werden als selbständige Wörter behandelt, solche mit einem einzigen Buchstaben dagegen werden mit dem folgender Wort zusammengeschrieben”

⁵⁸“läßt sich bei der Schreibung von Dialekttexten keine bestimmte Regel erkennen”

⁵⁹“Vor den Personen des Imperfekts außer der 1. Pers. sg. wird ħa- sehr selten mit Alif und häufig ohne Alif geschrieben. Ohne Alif kann ħa- entweder isoliert geschrieben oder mit dem imperfekt verbunden werden”

⁶⁰Les exemples donnés étant transcrits manuellement, il est difficile de voir s’il y a une espace entre les particules avec ’alif et le verbe qui suit. Dans le doute, nous en ferons figurer un.

⁶¹“ha- wird selten mit Alif geschrieben und meist mit dem Imperfekt verbunden.”

4.12.1.2.2 Particule proclitique /ma/

Farnawany (1981, p. 59) indique que la particule négative /ma/ s'écrit le plus souvent avec 'alif *ما قدمش* /ma ʔaddimʃ/ 'il n'a pas présenté', même devant une 'alif (ici de forme iv) *ما أمكنيش* /ma ʔamkanni:ʃ/ 'je n'ai pas pu'. Il ajoute que la 'alif de la particule peut être omise, quoique rarement : *مخلصناش* /ma xlisʕna:ʃ/ 'Nous n'avons pas fini'. Il précise que dans ce cas, la particule est toujours liée au mot qui la suit (voir aussi Farnawany (ibid., p. 167)). Dans tous ses exemples, /ʃ/, la partie enclitique de la négation, est aussi liée au mot.

Rosenbaum (2004, p. 301) cite des exemples d'emploi de la particule négative /ma/ écrite sans 'alif : *مشافوش* /maʃafu:ʃ/ 'ils n'ont pas vu' et *معرفش* /maʕrafʃ/ 'je ne sais pas'.

4.12.1.2.3 Préposition proclitique /ʕa/

Farnawany (1981, p. 59-60) indique que le /a/ de la préposition abrégée /ʕa/ 'sur' (de /ʕala/) s'écrit généralement avec 'alif, bien que la 'alif de l'article défini ne soit généralement pas écrite. Il donne trois graphies possibles : *دخلت عالباشحكيم* /daxalt ʕa l baʃhaki:m/ 'j'ai rencontré le médecin en chef' ('ayn lié à la 'alif de l'article), *قوم بنا عا الجامع* /ʔu:m bina ʕa l-ga:miʕ/ 'lève-toi, allons à la mosquée' ('ayn avec 'alif devant 'alif de l'article), *يتفرجوا ع الشقة* /jitfarragu ʕa ʃ ʃaʔʔa/ 'ils regardent l'appartement' ('ayn isolé devant 'alif de l'article). Farnawany (ibid., p. 170) cite la conjonction *عشان* /ʕaʃa:n/ ou *عشان* /ʕalaʃa:n/ 'parce que', issu de la fusion entre *على* /ʕala:/ et *شأن* /ʃaʔn/ 'sujet, question'.

Rosenbaum (2004, p. 304) illustre la transcription de /ʕa/ isolé sans 'alif par *ع الشارع* /ʕa ʃ-ʃa:riʕ/ 'dans la rue' et de /ʕa/ lié avec 'alif par *صبح عاليه* /sʕabbah ʕa l-be:h/ 'dis bonjour au bey'.

Avallone (2015) a observé dans ses échantillons les graphies suivantes de la préposition /ʕala:/ et de sa forme abrégée /ʕa/. Le nombre total d'occurrences étant indiqué après chacune d'elles : *على* 651, *علا* 85 (dont 80 dans *Rihla fī l-Nīl* (1965) 'Voyage sur le Nil' de 'Utmān Ṣabṛī), *ع* 108, *علا* 12, *علا* 5. La préférence pour la forme longue canonique est claire, mais la forme abrégée, essentiellement possible devant l'article défini, est tout de même présente dans la plupart des échantillons, lesquels couvrent plus d'un siècle de littérature en arabe égyptien. Pour la conjonction /ʕaʃa:n/ ou /ʕalaʃa:n/ 'parce que', les totaux sont les suivants : *عشان* 54 et *عشان* 25, attestés dans le corpus depuis *Rihla fī l-Nīl* (1965).

4.12.1.2.4 Prépositions proclitiques /fi/, /mi(n)/ et /bi/

Selon Farnawany (1981, p. 164), la préposition /fi/ 'dans' s'écrit avec *ف* *ف* isolé, rarement en prose et souvent en poésie. Il fournit l'exemple suivant : *ف بحر المينا* /fi baħr il-mina/ 'dans l'eau du port'. De même, il indique que la préposition /mi/ 'de' est écrite avec un *م* *م* isolé, fréquemment par certains écrivains et rarement par d'autres et il donne l'exemple *م الجهة دي* /mi-l-giha di/ 'de cette direction'. Farnawany (ibid., p. 170) mentionne une origine fuṣḥā des interrogatifs égyptiens *فين* /fe:n/ 'où?' < *في أين* /fi ʔajna/ et *منين* /mine:n/ 'd'où?' < *من أين* /min ʔajna/ pour lesquels les prépositions sont traitées en clitiques. Farnawany (ibid., p. 169) signale que la préposition /fi/ 'dans' peut être écrite

avec le nom suivant et donne l'exemple *عضها فيدها اليمن* /ʕadʕdʕaha f-idha l-jimi:n/ 'il lui mordit [dans] la main droite'.

Farnawany (1981, p. 162) ne relève qu'un exemple d'isolement de la préposition /bi/ (suivi de la *kašīda*) : *ب ه آلاف جنيه* : /bi xamast ala:f gine:/ 'de 5 mille livres', qu'il explique par le fait que la préposition est suivie d'un chiffre.

Rosenbaum (2004, p. 304) cite *مصر ف* /fi masʕr/ 'en Égypte', *نزل م القطر* /nizil mil-ʔatʕr/ 'il est descendu du train'.

De son côté, Avallone (2015) a observé dans ses échantillons les graphies suivantes pour /fi/ : 827 occurrences, *ف* 126, pour /mi(n)/ : 780, *م* 60 et pour /bi/ : 835, *ب* 13 (ces dernières uniquement dans *Muḏakkirāt tālib bi ʕa* 'Mémoire d'un étudiant boursier', écrit en 1942 par Lūwīs 'Awaḏ). Ici encore, la tendance prépondérante est au respect de la graphie *fuṣṣā*, mais les formes abrégées sont attestées et présentes dans de multiples échantillons pour *ف* depuis *Muḏakkirāt tālib bi ʕa* et pour *م* depuis *is-Sayyid wi mrātu f-maṣr* (1925) 'Monsieur et sa femme en Égypte' de Maḥmūd Bayram al-Tūnisī.

4.12.1.2.5 Préposition enclitique /li/

Pour les cas où la préposition /li/ associée aux pronoms personnels suffixés est suffixée au verbe pour exprimer un objet indirect, Farnawany (1981, p. 171) cite *جبتله* /gibtilu/ 'je lui ai apporté', *قولي* et *قولي* /ʔulli/ 'dis-moi' (impératif masculin) ainsi que *قويلو* /ʔuli:lu/ 'dis-lui' (impératif féminin).

Rosenbaum (2004, p. 301) cite notamment *افتحيله الباب* /ʔiftahi:lu l-ba:b/ 'ouvre-lui la porte' (impératif féminin) et *الكلام اللي بقولوك* /ʔil-kala:m illi balʔulhu:lak/ 'ce que je te [le] dis, ainsi que *قالي* et *قالي* /ʔalli/ 'il m'a dit'. Il cite aussi une occurrence de séparation entre verbe et pronom personnel affixe suivi de la préposition /li/ : *قدموا هالي* : /ʔaddimu ha:li/ 'donnez-la-moi' (avec emploi de al-'alif al-fāriqa (cf. page 38) placée après le wāw final de la forme plurielle du verbe) (ibid., p. 303).

Avallone (2015) relève 103 occurrences de verbes suivis de la préposition /li/ avec un pronom personnel, dont 62 sont isolées par une espace comme *قالي* /ʔalli/ 'il m'a dit', *أقول* /ʔaʔullak/ 'je te dis', et 41 sont liées au verbe, dont *قالي* ou *قالي* /ʔalli/ 'il m'a dit', *أقولك* /ʔaʔullak/ 'je te dis'. Par ailleurs, elle relève dans *is-Sayyid wi mrātu f-maṣr* (1925) 'Monsieur et sa femme en Égypte', de Maḥmūd Bayram al-Tūnisī, le mot *ماقولكيش* /maʔulki:f/ 'je te dis pas', d'une longueur de huit caractères, rendant son déchiffrement compliqué.

4.12.1.2.6 Préposition enclitique /bi/

Farnawany (1981, p. 170) cite comme exemple de la préposition /bi/ traitement enclitique de le verbe *جابه* /ga:b/ 'il a apporté' issu de *جاء ب* /ʕa:ʔa bi/ 'il est venu avec'.

4.12.1.2.7 L'article défini

Farnawany (ibid., p. 164) indique avoir trouvé deux occurrences de l'article défini écrit isolément : *ال ألف مليون فضان دول* /mi:n ḥayizraʕ lak il-ʔalf mi:ljo:n

fad^sd^a:n do:l/ 'qui cultivera pour toi ces mille millions de feddans⁶²?' et الإغراء /zihi?na mnil ?iğrā?/ 'nous en avons marre de l'incitation'⁶³. Dans les deux cas, l'article est placé devant un mot commençant par une 'alif marquant un "coup de glotte", ce qui peut contribuer à isoler graphiquement l'article, pour "casser" une trop longue suite de traits verticaux induits par 'alif et lām. Le fait est qu'en général l'article défini est presque exclusivement écrit en liaison avec le mot défini, comme en fuṣḥā.

4.12.1.2.8 Fusion de mots

4.12.1.2.8.1 Les déictiques sont régulièrement graphiés en un seul mot, comme chez Farnawany (1981) qui cite النهارده (ibid., [59]), النهاردة (ibid., [87]) et النهارضة (ibid., [156]) pour /ʔin-naharda/ 'aujourd'hui', دلوقت (ibid., [107]) et دلوقتي (ibid., [139]) pour /dilwa?ti/ 'maintenant'. Plus largement, Farnawany (ibid., [173]) cite des cas de fusion de noms et du démonstratif دي /di/ dans النواحيدي /innawahi:-di/ 'cette zone' et البلدي /il-ba'lad-di/ 'ce village', avec déplacement de l'accent tonique de /'balad/ par effet d'ajout du démonstratif, qui devient ainsi une partie du mot phonique, et donc logiquement du mot graphique.

Rosenbaum (2004, p. 304) cite des occurrences d'emploi du démonstratif دي /di/ avec الليلا دي /il-lēla:-di/ 'ce soir', السنادي /is-sana:-di/ 'cette année', ainsi que la fusion du démonstratif /da/ avec les pronoms personnels isolés /ʔana/ avec دنا ou دانا /dana/ 'c'est moi qui' et /ʔinta/ avec دنت ou دانت /dinta/ 'c'est toi qui'.

4.12.1.2.8.2 Les adjectifs numéraux cardinaux de 11 à 19 sont écrits en deux mots et un seul mot par Badawi et Hinds (1986) : أحاد عشر /hida:ʃar/ 'onze' (ibid., p. 9), اثنا عشر /ʔitna:ʃar/ 'douze' (ibid., p. 140), ثلاثة عشر /talatta:ʃar/ 'treize' (ibid., p. 134), أربعة عشر /ʔarbaʃta:ʃar/ 'quatorze' (ibid., p. 324), خمسة عشر /xamasta:ʃar/ 'quinze' (ibid., p. 266), ستة عشر /sitta:ʃar/ 'seize' (ibid., p. 398), سبعة عشر /sabaʃta:ʃar/ 'dix-sept' (ibid., p. 395), ثمانية عشر /tamanta:ʃar/ 'dix-huit' (ibid., p. 137), تسعة عشر /tisaʃta:ʃar/ 'dix-neuf' (ibid., p. 129).

Boutros (2000, p. 86) emploie les mêmes graphies en un seul mot, sauf pour ثنتاشر /tamanta:ʃar/ 'dix-huit', dont il supprime la 'alif ne correspondant pas à un /a:/ long.

Farnawany (1981) cite uniquement des occurrences de graphies en un seul mot avec حداشر /hida:ʃar/ 'onze' (ibid., p. 107), خمستاشر /xamasta:ʃar/ 'quinze' (ibid., p. 171), ستاشر /sitta:ʃar/ 'seize' (ibid., p. 77), اثناشر /ʔitna:ʃar/ 'douze' (ibid., p. 51), اثناشر /ʔitna:ʃar/ 'douze' (ibid., p. 51), تسعطاشر /tisaʃta:ʃar/ 'dix-neuf' (ibid., p. 103) (avec pour ces deux derniers pharyngalisation de /t/ transcrite par le graphe ʔā ط au lieu de tā ت).

Rosenbaum (2004, p. 310) indique que ces fusions sont prononcées et habituellement orthographiées en un seul mot, comme حداشر باب /hida:ʃar ba:b/ 'onze portes', اثناشر لوري /ʔitna:ʃar lo:ri/ 'douze camions', سبعة عشر مواطن /sabaʃta:ʃar muwa:tʃin/ 'dix-sept citoyens' ainsi que كام سنة ؟ خمستاشر ؟ ستاشر ؟ /ka:m sana ? xamasta:ʃar ? sitta:ʃar ?/ 'Combien d'années ? Quinze ? Seize ?'. Il ajoute que la graphie en deux mots est parfois (pas souvent)

⁶²Mesure agraire de référence en Égypte qui correspond à une superficie de 60 mètres sur 70, soit 4.200 m² (0,42 hectare).

⁶³Dans le deuxième exemple, est reproduite la manière hésitante de parler d'une personne (l'article est répété), de sorte que, selon l'auteur, il est difficile de parler d'écriture isolée.

attestée et cite تاسع /tisaʕ ta:ʕar/ 'dix-neuf', avec le tā' ت préfixé à /a:ʕar/ et non au mot précédent, bien qu'il soit issu d'une tā' marbūṭa suffixée à ce dernier. Il relève une graphie exceptionnelle en deux mots avec خمسة اشر يوم /xamasta:ʕar jo:m/ 'quinze jours', avec conservation de la tā' marbūṭa étymologique.

4.12.1.2.8.3 Les centaines sont écrites en un seul mot par Badawi et Hinds (1986) avec la terminaison étymologique مائة /tuṭumijja/ 'trois-cents' (ibid., p. 134), ربيعمة et أربعمة /rubʕumijja/ 'quatre-cents' (ibid., p. 324), خمسمائة /xumsumijja/ 'cinq-cents' (ibid., p. 266), ستمة et ستمية /sutumijja/ 'six-cents' (ibid., p. 398), سبعمة et سبعمة /subʕumijja/ 'sept-cents' (ibid., p. 395), ثمانمة et تمنمة /tumnumijja/ 'huit-cents' (ibid., p. 137), تسعمائة /tusʕumijja/ 'neuf-cents' (ibid., p. 129).

Farnawany (1981) cite des occurrences de noms de centaines écrits en un seul mot, en situation d'annexion, avec la terminaison /mi:t/ : تلتمة /tultumi:t bayya:ʕ/ 'trois-cents vendeurs' (ibid., p. 79), ستمية /sutumit kabtin/ 'six-cents capitaines' (ibid., p. 163).

Rosenbaum (2004, p. 309) indique que les centaines de 300 à 900 sont habituellement écrites en un seul mot et cite les formes isolées خمسمية /xumsumijja/ 'cinq-cents' et ستمة /sutumijja/ 'six-cents' ainsi que les formes en annexion تلتمة /tultu'mi:t gine:/ 'trois-cents livres' et ستمية /xumsu'mi:t suttu'mi:t ʕaʔl/ 'cinq-cents six-cents cerveaux'.

4.12.1.2.8.4 Les milliers sont plus rarement écrits en un seul mot selon Rosenbaum (ibid., p. 309) qui cite سبعة آلاف /sabʕat ala:f tʕann/ 'sept mille tonnes'. De son côté, Farnawany (1981, p. 171) cite خمسة آلاف /xamas tala:f/ 'cinq mille'.

De manière originale, Girod et Aziz (2016, p. 92) transcrivent le /t/ liant les unités et /ala:f/ par un tā' suivi de la kašida : ثلاث ت الاف /talat t ala:f/ 'trois mille', أربع ت الاف /ʔarbaʕ t ala:f/ 'quatre mille', خمس ت الاف /xamas t ala:f/ 'cinq mille', ست ت الاف /sitt ala:f/ 'six mille'.

Rosenbaum (2004, p. 310) relève des occurrences de graphies en deux mots avec خمس عشرة آلاف /xamas tala:f gine:/ 'cinq mille livres' et عشرة آلاف /ʕaʕar tala:f gine:/ 'dix mille livres', le /t/ étant lié à /ala:f/ par un tā' ت formant le néologisme تلاف /tala:f/.

4.12.2 Orientations Wikipédia maṣri

4.12.2.1 Recommandations officielles

Les recommandations qui figurent sur la page officielle طريقة الكتابة Tārīqa al-kitāba 'méthode d'écriture' (Wikipédia, 2025c) donnent les instructions suivantes :

ادوات الجر و ادوات العطف
كتابة ادوات الجر و ادوات العطف اوضح و اسهل لو تكتب مفصوله عن الكلمه الى بتيجي وراها. مثلا: <في اوروبا> او <ف اوروبا> اوضح و اسهل للقرايه من <فاوروبا>.

Prépositions et conjonctions

L'écriture des prépositions et des conjonctions est plus claire et plus facile si

elles sont écrites séparées du mot qui les suit.” Par exemple : ف في اوروبا ou ف في اوروبا sont plus clairs et plus faciles à lire que فا اوروبا /fi ʔ urubba/ ‘en Europe’ (Ghaly, le 10/05/2008).

La séparation des prépositions et conjonctions est donc la recommandation officielle. Plus généralement, la question des éléments à affixer ou non apparaît comme un enjeu, notamment dans la page de discussion (archive 1), où un contributeur anonyme s’interroge sur la manière d’écrire à retenir : عملت له ou عملته /ʕamalti lu/ ‘je lui ai fait’, avec affixation de la préposition suivi d’un pronom, ou non affixation, comme en fuṣṣā. Sur la même page, on trouve aussi بعثنا contre ه ل نا بعث /baʕatthalna/ ‘tu nous l’as envoyée’ ou ماديهاوش contre ما اديت ها لوش /ma ʔiddithalu:f/ ‘je ne la lui ai pas donnée’. De son côté, Maysara (archive 2) écrit spontanément le 24/10/2009 l’expression ما-عندنا-ش /maʕandina:f/ ‘nous n’avons pas’ où elle sépare par des tirets les deux éléments de la négation de part et d’autre de la préposition suivie du pronom affixe. Ces options graphiques différentes illustrent le dilemme entre une écriture jointe qui rend compte de l’unicité du mot phonique et fluidifie le geste du scripteur et l’utilisation d’espaces ou de tirets pour clarifier la composition du mot et faciliter la tâche du lecteur.

Un contributeur anonyme fait référence au livre *Hādir al-ṭaqāfa fī miṣr* حاضر الثقافة في مصر “Le présent de la culture en Égypte” (1999) de Bayūmī Qandīl (1942-2009), un linguiste égyptien. Dans ce livre celui-ci appelle à ne pas affixer en ‘āmmyya, y compris les pronoms personnels, pourtant affixés en fuṣṣā : ح تفعل مخ /ḥatiʔfil muxxak/ ‘ton cerveau va s’arrêter’ et ح نطخ ك /ḥantʕuxxak/ ‘on va te frapper’ (Qandīl, 1999, p. 182).

4.12.2.2 Traitement opéré par Ghalybot

Pour les prépositions, Ghalybot n’intervient pas pour les séparer des mots suivants. Il remplace ف تقع par ف ‘situé dans’, entérinant un usage abrégé de la préposition في /fi/, mais il garde le mot complet quand il remplace عمومًا par في العموم ‘en général’. Pour من منذ il conserve la forme fuṣṣā non abrégée en remplaçant من منذ par من منذ ‘depuis’ et على ايد par على ايد ‘en présence de’. L’abrégement ayant lieu avant une suite de deux consonnes, comme celle formée par l’article défini et le début du mot suivant, son traitement automatisé a probablement été ressenti comme non discriminant. Dans la suite espace ل espace, Ghalybot efface la deuxième espace, ce qui indique qu’il souhaite que la préposition /li/ soit liée au mot qui la suit.

Ghalybot ajoute une espace après و /wi/ ‘et’ dans de nombreuses instructions, dont 1 و (et tous les chiffres jusqu’à 9), ainsi que dans كانت و /wi ka:nit/ ‘et elle était’, notamment. Ce choix est à rebours de la pratique en fuṣṣā, où cette conjonction est cliticisée au mot qui suit. Outre le respect des recommandations officielles, il semble y avoir ici un désir de distinguer l’écriture de l’arabe égyptien par un détail, qui est cependant très récurrent.

Ghalybot remplace la particule fuṣṣā du futur سوف /sawfa/, précédée et suivie d’une espace, par la particule ‘āmmyya ها /ha/, seulement précédée d’une espace, ce qui implique qu’il souhaite qu’elle soit liée au verbe qui suit. Une instruction ultérieure est similaire sauf qu’une espace est ajoutée après /ha/, contredisant son premier choix. Dans sept instructions, il colle la particule au verbe, en remplaçant, par exemple, هايشكل سيشكل par هايشكل ‘il formera’, mais dans six autres instructions, il place une espace entre la particule et le

verbe en remplaçant, par exemple, *سيفض* par *ها يرفض* ‘il refusera’, sans qu’une logique soit décelable dans ces choix contradictoires.

La même pluralité des options est observée dans le traitement de la particule *ما* /ma/ de la négation avec l’absence d’espace la séparant du verbe suivant dans 23 instructions, par exemple quand il remplace *لم يكن* par *ما كانش* ‘il n’était pas’. Dans 14 autres instructions, il place une espace entre particule et verbe, notamment quand il remplace *لم تكن* par *ما كانتش* ‘elle n’était pas’.

Pour les adjectifs numériques de 11 à 19, Ghalybot substitue les formes ordinales ‘ām-miyya en un seul mot à celles en fuṣḥā en deux mots. Ainsi, il remplace *الحادي عشر* par *الحادشر* ‘le onzième’, *الثاني عشر* par *الاتاشر* ‘le douzième’, *الثالث عشر* par *التلاتاشر* ‘le treizième’, *الرابع عشر* par *الاربعتاشر* ‘le quatorzième’, *الخامس عشر* par *الخمستاشر* ‘le quinzième’, *السادس عشر* par *الستاشر* ‘le seizième’, *السابع عشر* par *السبعتاشر* ‘le dix-septième’, *الثامن عشر* par *التمنتاشر* ‘le dix-huitième’, *التاسع عشر* par *التسعتاشر* ‘le dix-neuvième’.

4.12.3 Observations dans le corpus

4.12.3.1 Exemples de choix graphiques

4.12.3.1.1 Particules proclitiques du futur

Dans Wikipédia maṣri, les particules du futur sont transcrites selon les graphies déjà décrites. Quand elles sont séparées du verbe qui suit, /ha/ peut être transcrit par *ح* isolé ou avec ’alif, tandis que /ha/ est généralement transcrit par *ه* avec ’alif, comme l’indique le tableau 4.64, qui recense les occurrences de ces graphies suivies des préfixes *ي* de la troisième personne singulier masculin ou pluriel et *ت* de la troisième personne singulier féminin ou deuxième personne singulier ou pluriel de l’inaccompli subjonctif qui suit ces particules.

API	ح			ها			total
		#	%		#	%	
/ħaj.../	ح ي	55	16	حاي	304	84	359
/ħat.../	ح ت	58	41	حات	85	59	143
API	ه			ها			total
		#	%		#	%	
/ħaj.../	ه ي	4	04	هاي	97	96	101
/ħat.../	ه ت	0	0	هات	73	100	73
total		117	18		559	82	676

Tab. 4.64: Particules du futur *ح* et *ها* séparées du verbe

Il a ainsi été trouvé deux occurrences de *ها يخذ* /ha jaxud/ ‘il prendra’ (dont une dans *شكلك مش غريب*), une occurrence de *ها يعني* /ha jaʕni/ ‘ça signifiera’ (Second rôle) et une occurrence de *ها يعتبر* /ha juʕtabar/ ‘ça sera considéré’ (Juziers, commune des Yvelines).

Pour les particules graphiquement liées au verbe suivant, le tableau 4.65 récapitule le nombre d’occurrences pour le verbe très commun correspondant à ‘être’.

traduction	API	ح			حا			total
			#	%		#	%	
il sera	/hajku:n/	حيكون	38	50	حايكون	39	50	77
elle sera	/hatku:n/	حتكون	23	58	حاتكون	17	42	40
traduction	API	هـ			ها			total
			#	%		#	%	
il sera	/hajku:n/	هيكون	164	35	هايكون	316	65	480
elle sera	/hatku:n/	هتكون	62	39	هاتكون	100	61	162
total			287	38		472	62	759

Tab. 4.65: Particules du futur ح et حا liées au verbe

La situation est contrastée pour /ha/, mais pour /ha/, la préférence est nettement à son emploi avec la 'alif, sans espace avec le verbe suivant. هاتكون est le choix de Ghalybot pour remplacer l'équivalent fuṣḥā ستكون, alors que c'est ها يكون qu'il choisit pour remplacer سيكون. Ces remplacements ne semblent pas imposer un choix graphique hégémonique.

4.12.3.1.2 Particule proclitique /ma/

Pour la particule /ma/ de négation (complétée par l'enclitique /f/ régulièrement écrit ش), les graphies possibles comprennent mīm م lié directement au mot suivant, avec 'alif collé au mot suivant et avec 'alif séparé du mot suivant par une espace. Le tableau 4.66 récapitule différents emplois courants de cette particule.

traduction	API	م			ما			ما-			total
			#	%		#	%		#	%	
personne	/maḥaddi:f/	محدث	129	41	ماحدث	106	33	ما حدث	83	26	318
il n'y a pas	/mafi:f/	مفيش	470	67	مافيش	168	24	ما فيش	64	09	702
elle n'a pas	/malha:f/	ملهاش	8	04	مالهاش	220	87	ما لهاش	23	09	251
il n'était pas	/makan:f/	مكانش	98	06	ماكانش	1.297	75	ما كانش	341	19	1.736
total			705	24		1791	60		511	16	3007

Tab. 4.66: Graphies de la particule négative /ma ... f/

La graphie composée de mīm م avec 'alif collé au mot suivant est prépondérante, même si pour /maḥaddi:f/ et /mafi:f/ c'est mīm م lié directement au mot suivant qui l'emporte, malgré le remplacement de لا يوجد par مافيش effectué par Ghalybot. Par contre, le choix de ce dernier de remplacer لم يكن par ماكانش contribue probablement à la popularité de cette graphie, même si plus loin dans le code du "bot", "لم يكن هناك" est remplacé par ماكانش فيه 'il n'y avait pas'. La présence ou l'absence de la 'alif ou d'une espace semble aléatoire, l'important étant que ce soit déchiffable par le lecteur.

Cette double particule, notamment cliticisée à des verbes de forme x, induit des chaînes de caractères particulièrement longues pour l'arabe, avec 11 caractères pour مايبستعملوش /majistaxdimuha:f/ 'ils ne l'utilisent pas' (كلمة) Fuck et pour مايبستعملوش /majbistaʕmlu:f/ 'ils n'emploient pas' (بولاندى) Polonais). Cette longueur inhabituelle

complexifie le déchiffrement de la chaîne écrite par le lecteur. Des suites de caractères plus courtes peuvent aussi paraître obscures, notamment quand les lettres de la racine sont noyées dans une masse d'affixes et de clitiques, comme dans cette séquence de huit caractères : ماقلتلكش /maʔultilakʃ/ 'je ne t'ai pas dit' (ماقلتلكش ان الكثرة).

4.12.3.1.3 Préposition proclitique /ʕa/

La préposition /ʕa/ placée devant l'article défini est écrite de manière dominante comme en fuṣḥā على al avec 68.927 occurrences. On trouve aussi une 'alif simple à la place de la 'alif maqṣūra avec 249 occurrences de علا. Quand la préposition est abrégée en /ʕa/ devant l'article, les graphies possibles comprennent 'ayn ع isolé, avec 'alif collée à la 'alif de l'article et avec 'alif séparée de la 'alif de l'article par une espace, comme illustré dans le tableau 4.67.

traduction	API	ع			عا			عاـ			total
			#	%		#	%		#	%	
sur la/le(s)	/ʕa-l.../	ع ال	1.485	99	عال	4	0,005	عا ال	15	0	1.504

Tab. 4.67: Graphies de la préposition abrégée /ʕa-l...

La graphie la plus courante de /ʕa-l.../ est donc celle de 'ayn ع isolé séparé de l'article par une espace, comme par exemple pour le titre de l'article consacré à la chanson ع الخطه /ʕa-lmaḥatʔa/ 'à la gare'. La graphie عال se trouve notamment dans ع الطريق /ʕa-tʔ-ʔari:ʔ/ 'sur la route', dans le texte de la chanson وانا على الربابه باغنى 'et moi sur le rabâb⁶⁴ je chante'. La graphie عا se trouve par exemple dans عا الجبل /ʕa-l-gabal/ 'sur la montagne', dans le texte de la chanson الجيش رجع م الحرب 'L'armée est revenue d'la guerre'. L'emploi de cette préposition abrégée semble donc particulièrement utile pour adapter la prosodie de la langue au rythme des chansons.

Il a été dénombré 8.085 occurrences de la conjonction عشان /ʕaʃa:n/ et 8.447 occurrences de sa forme abrégée عشان /ʕaʃa:n/ 'parce que'.

4.12.3.1.4 Prépositions proclitiques /fi/, /mi(n)/, /bi/ et /li/

4.12.3.1.4.1 La préposition /fi/ placée devant l'article défini est écrite de manière dominante en deux graphes soit avec 'alif maqṣūra⁶⁵ في al avec 306.451 occurrences, ou comme en fuṣḥā في al avec 5.918 occurrences. On dénombre seulement 3.565 occurrences de la graphie en un graphe fā' isolé par une espace ف ال, comme dans الشمال /fi-f-ʃima:l/ 'dans le nord' dans l'article جنوب السودان Soudan du Sud. Il n'a pas été trouvé d'occurrences du graphe fā' lié à la 'alif de l'article فال avec le sens de la préposition /fi/ 'dans', mais on trouve cette suite dans l'article cité ci-dessus pour transcrire la conjonction /fa/ comme dans فالتقديرات النسبية /fa-l-taʔdira:t in-nisbija/ 'alors les estimations relatives', sans différenciation d'avec la fuṣḥā.

⁶⁴Rabâb est le nom générique des violons à archet dans le monde arabe.

⁶⁵Selon la graphie du yā' final sans point promue par Wikipédia maʃri, et Ghalybot, comme étant celle pratiquée en Égypte.

4.12.3.1.4.2 La préposition /mi(n)/ placée devant l'article défini est écrite de manière dominante en deux graphes comme en fuṣḥā من ال avec 916.936 occurrences recensées. On dénombre seulement 647 occurrences de la graphie en un graphe mīm isolé par une espace م ال, comme dans انواع م الاشعاع /ʔanwa:ʕ mi-l-ʔiʕʕa:ʕ/ 'des types de radiations' dans l'article فيزيا نوويه Physique nucléaire. Il n'a pas été trouvé d'occurrences du graphe mīm lié à la 'alif de l'article مال avec le sens de la préposition /mi(n)/ 'de',

4.12.3.1.4.3 La préposition /bi/ placée devant l'article défini est écrite de manière dominante liée comme en fuṣḥā بال par exemple dans بالحروف /bil-huru:f/ 'avec les lettres' (141 occurrences). On dénombre seulement 159 occurrences de la graphie en un graphe bā isolé par une espace ب ال, dont 16 occurrences de ب الحروف comme dans يكتبوها ب بيكتوبوها ب /bijiktibu:ha bil-huru:f il-lati:ni/ 'ils l'écrivent avec les lettres latines' dans l'article مصر مصر Égypte. On trouve une autre occurrence dans المزيكه مهتمه ب /mihtamma bi-l-mazzi:ka/ 'intéressée par la musique' dans l'article سميره موسى Samira Moussa⁶⁶.

4.12.3.1.4.4 La préposition /li/ a tendance à être séparée graphiquement du mot qui la suit par certains rédacteurs d'articles, comme Masrialltheway dans l'article سميره موسى Samira Moussa. En novembre 2021, il ajoute plusieurs paragraphes avec 12 emplois de lām isolé dont التطبيقات الفيزيا النوويه ل لتطبيق السلمى ل لاستعمال السلمي /bitidʕu lil-istiʕma:l is-silmi li-tatʕbiʔa:t il-fizja n-nawawijja/ 'elle appelle à l'utilisation pacifique des applications de la physique nucléaire'. Le premier lām isolé est placé avant un lām de l'article (sans 'alif) introduisant un nom commençant par une 'alif. La règle fuṣḥā d'effacement de la 'alif de l'article après /li/ est respectée, sauf pour ce qui est de l'espace ajoutée. Le deuxième lām isolé est placé devant un nom sans article. La plupart des autres lām isolés de cet article sont placés devant un article défini complet comme dans وانتقلوا ل القاهرة /wi tnaʔalu lil-qa:hira/ 'et ils ont déménagé au Caire', où l'espace et la 'alif de l'article contreviennent à la règle fuṣḥā. Ces espaces seront effacées par Ghalybot en mai 2023.

4.12.3.1.5 Préposition enclitique /li/

Le traitement de la suite préposition /li/ + pronom personnel après un verbe (avec valeur d'objet indirect) s'oriente sur Wikipédia maṣri vers une cliticisation de ces éléments, à rebours de la règle fuṣḥā, comme l'illustre le tableau 4.68.

Une très forte tendance se dessine en faveur de la cliticisation, avec les formes majoritaires suivantes :

- قالهم /ʔalluhum/ 'il leur a dit' شهدا كوردوبا Martyrs de Cordoue), l'absence d'occurrences de قالهم semble indiquer une préférence à la simplification de l'écriture des deux lām ل en les fusionnant, même si l'un fait partie du verbe, et l'autre de la préposition. La gémination du lām n'est jamais explicitée par une šadda ّ comme dans قالهم.
- اداله /ʔiddalu/ 'il lui a donné' موليير Molière). L'hégémonie de la graphie liée semble indiquer ici un respect de la règle fuṣḥā de transformation de la 'alif maqṣūra du

⁶⁶Samira Moussa (1917-1952) était une scientifique égyptienne, pionnière dans la recherche nucléaire.

traduction	API	lié			séparé			total
			#	%		#	%	
il leur a dit	/ʔalluhum/	قالهم	81	88	قال لهم	12	12	93
		قاللهم	0	0				
il lui a donné	/ʔiddalu/	اداله	63	93	اداله	3	4	68
					أدى له	2	3	
il lui vient	/jigi:lu/	يجيله	37	90	يجي له	2	5	41
					يجي له	2	5	
il lui a apporté	/gablu/	جابله	9	70	جاب له	4	30	13
total			190	89		25	11	215

Tab. 4.68: Graphies de la préposition /li/ + pronom personnel

verbe seul en 'alif simple devant un suffixe cliticisé.

- يجيله /jigi:lu/ 'il lui vient' (ديسلكسيا Dyslexie), pour cette occurrence précise, le rédacteur a précisé la prononciation du pronom suffixe par une ḍamma يجيله.
- جابله /gablu/ 'il lui a apporté' (حادثة المنصبه Assassinateur d'Anouar el-Sadate), l'origine prépositionnelle du /b/ de /ga:b/ disparaît complètement avec l'ajout de la préposition /lu/.

4.12.3.1.6 L'article défini

Dans l'encyclopédie, plusieurs occurrences de l'article défini /(ʔi)/ sont transcrites par la suite de graphes 'alif + lām ال séparée du mot suivant par une espace, à rebours de la règle fuṣḥā :

- Devant un nombre écrit en chiffres avec une kaṣīda comme dans الولايات ال 48 /ʔil-wilaja:t it-tamanja wi ʔarbaʕi:n il-mutgawra/ 'les 48 états contigus' (المتجاورة Barack Obama) ou sans kaṣīda comme dans بلد ال 240 في العالم /ʔil-mite:n wi ʔarbaʕi:n balad illi fi l-ʕa:lam/ 'les 240 pays qu'il y a dans le monde' (بي BBC News).
- Devant des lettres latines, par exemple un mot comme dans Unicode ال 'Unicode' (Wikipédia : problèmes de police) ou par exemple un sigle comme dans DNA ال /fi l-di ʔin ʔe:/ 'dans l'ADN' (سرطان Cancer). Le terme arabe existe pourtant puisque l'article correspond à ADN porte le titre حمض نووي.
- Devant un nom souvent d'origine étrangère commençant par une 'alif sans hamza comme dans باراك أوباما /fi l-ʔo:dʕa l-xadʕra/ 'dans le salon vert' (Barack Obama) ou avec hamza comme dans بلديات ال ألامو /baladijjat il-ʔalamu wi.../ 'les communes d'El Álamo et...' (مطار مدريد-باراخاس Aéroport Adolfo Suárez Madrid-Barajas)

- Sans raison apparente sans assimilation de /l/ comme dans *ال غلط* /ʔil-ʔalat/ 'l'erre-
reur' (إعدام Peine de mort) ou avec assimilation de /l/ comme dans *ال ارتفاع* /ʔil-irtifaːʕ id-dimugraːfi/ 'la hausse démographique' (روما Rome).

4.12.3.1.7 Fusion de mots

4.12.3.1.7.1 Les adjectifs numériques de 11 à 19 sont écrits de façon prépondérante en un seul mot, selon la prononciation 'āmmiyya et les graphies fuṣḥā ou proches sont très minoritaires, comme l'illustre le tableau 4.69.

nombres	API	fuṣḥā			'āmmiyya			total
			#	%		#	%	
onzième	/hidaːʃar/	حادى عشر	177	17	حداشر	910	83	1087
douzième	/ʔitnaːʃar/	ثانى عشر	13	02	اتناشر	976	98	989
treizième	/talattaːʃar/	تالت عشر	31	04	تلاتاشر	868	96	899
quatorzième	/ʔarbaʕtaːʃar/	رابع عشر	157	19	اربعتاشر	711	81	868
quinzième	/xamastaːʃar/	خامس عشر	97	13	خمستاشر	687	87	784
seizième	/sittaːʃar/	سادس عشر	102	09	ستتاشر	1104	91	1206
dix-septième	/sabaʕtaːʃar/	سابع عشر	32	04	سبعتاشر	775	96	807
dix-huitième	/tamantaːʃar/	ثامن عشر	36	04	تمنتاشر	877	96	913
dix-neuvième	/tisaʕtaːʃar/	تاسع عشر	151	15	تسعتاشر	868	85	1019
total			796	10		7776	90	8572

Tab. 4.69: Graphies des adjectifs numériques de 11 à 19

Les formes ordinales de ces dernières (choisies pour le tableau) sont plus nombreuses que les formes cardinales, probablement de par leur utilisation pour la numérotation des siècles. Elles sont pourtant remplacées par Ghalybot par leurs équivalents 'āmmiyya qui correspondent aux cardinaux et aux ordinaux, remplacement donc limité. La forme fuṣḥā cardinale la plus courante dans le corpus est اثنا عشر 'douze' avec 160 occurrences, les autres ne représentant que quelques dizaines, voire quelques unités.

4.12.3.1.7.2 Les centaines dans Wikipédia maṣri sont très rarement écrites en deux mots. Quand elles sont écrites en un seul mot, la tendance est à conserver la graphie étymologique, ce qui justifie les choix graphiques de Badawi et Hinds (1986). Les graphies qui transcrivent la prononciation réelles en arabe égyptien sont attestées, mais minoritaires, comme l'illustre le tableau 4.70. Il faut noter le nombre relativement réduit d'occurrences⁶⁷, la tendance étant d'écrire les centaines en chiffres.

4.12.3.1.7.3 Les milliers sont peu écrits en lettres et quand ils le sont, ils sont le plus souvent écrits en deux mots selon la graphie fuṣḥā. On trouve aussi des occurrences de graphies en un seul mot, qui reflètent plus la prononciation 'āmmiyya, comme l'illustre le tableau 4.71.

⁶⁷ Il n'a été trouvé que quelques unités pour chaque forme en isolée en حمية, le maximum étant cinq pour خمسمية /xumsumijja/ 'cinq-cents'. Ces formes n'ont pas été indiquées dans le tableau.

Cent-aines	API	fuṣḥā deux mots			fuṣḥā un mot			‘āmmiyya un mot			total
			#	%		#	%		#	%	
300	/tultu'mi:t/	ثلاث مائة	2	07	ثلاثمائة	16	56	تلتميت	11	37	29
400	/rubʕu'mi:t/	أربع مائة	0	0	أربعمائة	12	55	ربعميت	10	45	22
500	/xumsu'mi:t/	خمس مئة	1	02	خمسمئة	34	68	خسميت	15	30	50
600	/suttu'mi:t/	ست مئة	1	05	ستمئة	9	45	ستيت	10	50	20
700	/subʕu'mi:t/	سبع مائة	2	15	سبعمئة	6	43	سبعميت	6	42	14
800	/tumnu'mi:t/	ثمان مائة	0	0	ثمانمائة	6	86	تمنيت	1	14	7
900	/tusʕu'mi:t/	تسع مائة	0	0	تسعمائة	7	88	تسعمية	1	12	8
total			6	04		90	60		54	36	150

Tab. 4.70: Graphies des centaines

Milliers	API ‘āmmiyya	fuṣḥā			‘āmmiyya			total
			#	%		#	%	
3.000	/talatt ala:f/	ثلاثة آلاف	4	80	تلتلاف	1	20	5
4.000	/ʔarbaʕt ala:f/	أربعة آلاف	5	84	اربعتلاف	1	16	6
5.000	/xamast ala:f/	خمسة آلاف	13	77	خمستلاف	4	23	17
6.000	/sitt ala:f/	سنة آلاف	2	40	ستلاف	3	60	5
7.000	/sabʕat ala:f/	سبعة آلاف	5	100	سبعتلاف	0	0	5
8.000	/tamant ala:f/	ثمانية آلاف	4	67	تمنتلاف	2	33	6
9.000	/tisaʕt ala:f/	تسعة آلاف	1	100	تسعتلاف	0	0	1
10.000	/ʕaʕart ala:f/	عشرة آلاف	20	52	عشرتلاف	19	48	39
total			54	20		30	35	84

Tab. 4.71: Graphies des milliers

Parmi les occurrences de graphies en un seul mot, on trouve : ثلاث مائت 'trois mille vers' et اربعتلاف غلبان (Abdullah an-Nadeem), 'quatre mille vers' (النديم), 'cinq mille pauvres' (سيرة الظاهر بيبرس), 'environ six mille ans avant J.C.' (Période prédynastique égyptienne), 'huit mille mètres' (Liste des plus hauts sommets sur Terre), 'Dix mille combattants' (Baybars al-Mansuri), 'une légion composée de six cent soixante-six soldats' (Massacre de la légion thébaine).

On trouve également quelques occurrences de graphies en deux mots visant à transcrire la prononciation ‘āmmiyya. L'enjeu est alors la place de /t/ issu de la tā' marbūṭa. Une seule occurrence a été trouvée où il est lié au premier mot, sa place d'origine en fuṣḥā avec عشرتلاف 'dix mille dinars' (الصالح ايوب) Malik al-Salih Ayyoub). Les autres occurrences lient ce /t/ à /ala:f/, dont la 'alif est effacée : ثلاث مائت 'trois mille ans' (الفن), 'cinq mille soldats' (ثيودورا) 'pour sept mille' (Theodora Cantacuzène (épouse d'Orhan)),

ans’ حسن البنا) ‘de dix mille livres’ بعشر تلاف جنيهه), (الخليج الفارسي) Hassan el-Banna). On retrouve ici le traitement graphique de la prononciation de la terminaison /t/ des chiffres de trois à neuf devant une ’alif effacée.

4.13 Conclusion

Wikipédia maṣri se révèle ainsi comme un riche recueil des nombreuses graphies de l’arabe égyptien, qu’elles aient déjà été décrites ou non. Les recommandations “orthographiques” de la page officielle sont souvent appliquées, que ce soit par Ghalybot, le programme massivement utilisé pour “égyptianiser” les textes des articles, mais aussi par les rédacteurs ordinaires, consciemment ou inconsciemment. De multiples expérimentations graphiques contribuent à expliciter la phonologie de cette variété d’arabe. Les graphes ’alif ʾ, wāw و et yā ʾي, voire ’alif maqṣūra ى sont régulièrement employés pour matérialiser les voyelles courtes, orientant le système d’écriture vers une tendance phonographiques, et l’éloignant de son utilisation en tant qu’abjad. Les graphes alternatifs پ et ف, ج sont employés sans complexe pour transcrire les sons des emprunts étrangers. La cliticisation est aussi appliquée avec largesse pour prendre en compte la réalité spécifique du mot phonique ʾāmmiyya. Ces choix hétérodoxes d’écriture apparaissent souvent comme un affichage revendiqué de la différence entre arabe égyptien et arabe littéral, conférant ainsi une dimension politique à ces options graphiques. Malgré tout, l’orthographe fuṣḥā est souvent respectée, faute d’alternative parfois, mais aussi souvent par habitude. La langue écrite reste manifestement de l’arabe, malgré les dénégations de certains contributeurs. Il reste que toutes les graphies “autres” testées ouvrent des perspectives d’écritures dialectales légitimables dans le futur.

Conclusion

L'écriture de l'arabe égyptien sur Wikipédia maṣri est une expérience unique en son genre. La rédaction d'articles encyclopédiques dans une langue non normée, et dont l'usage relève essentiellement de la vie quotidienne, à l'origine rurale, s'avère un pari très partiellement réussi. L'enthousiasme des premiers rédacteurs, louable face aux nombreuses critiques qui leur ont été adressées, a permis la publication d'articles, souvent en lien avec des enjeux nationaux, des personnalités égyptiennes et autres réalités locales, tout ceci confirmant la dimension nationaliste du projet. Peu à peu, le désir d'être classé parmi les premières éditions Wikipédia en terme de nombre d'articles a amené à privilégier la quantité par rapport à la qualité, notamment par la rédaction automatisée d'articles à partir de la base de donnée Wikidata, sans grande valeur ajoutée en comparaison de celle-ci. Face aux critiques pointant du doigt le grand nombre d'articles au stade d'ébauche, l'effort s'est plus récemment réorienté vers un étoffement de ces derniers, mais cette fois encore par des moyens artificiels. Il y a tout d'abord la copie à peine modifiée d'articles déjà publiés sur Wikipédia al-ʿarabiyya, et plus subtilement la traduction d'articles publiés dans l'édition en anglais. Le résultat est une langue arabe de variété fuṣḥā, la seule que puisse produire un traducteur automatique, qui est ensuite traitée par un "bot" rudimentaire pour "égyptianiser" cette langue, avec les résultats mitigés qu'on peut attendre d'un tel dispositif. Bref, l'intérêt de l'ensemble du projet reste limité, puisque le public potentiellement intéressé lit au minimum l'arabe, et donc peut consulter la version al-ʿarabiyya, bénéficiant d'une plus grande profondeur d'articles et aussi d'un plus grand prestige.

Les articles de cette encyclopédie en arabe égyptien constituent pour autant un corpus pertinent pour l'étude de l'écriture de l'arabe égyptien. Nombre d'entre eux sont des compositions dans une langue authentique et naturelle. Il a ainsi été possible de mettre en évidence des stratégies d'écriture individuelles (transcription du "coup de glotte" issu de qāf par hamza, rattachement de tā' au nom compté, ...). Des usages partagés ont aussi pu être mis en évidence par les relevés statistiques (usage décomplexé des graphes ج, پ et ف, différenciation de la tā' marbūṭa prononcée /a/ sans points de celle prononcée /it/ avec points, ...).

Cette expérience singulière est d'ailleurs une formidable occasion d'être témoin d'une réflexion sur leur langue par des locuteurs natifs pour la plupart non spécialistes. Les discussions préalables au projet ainsi que la rédaction de la page des recommandations exposent la perception que ceux-ci ont de l'enjeu de la mise à l'écrit d'une langue non normée. Ils s'inscrivent dans la tradition des dictionnaires bilingues de Socrates Spiro, à la fin du xix^e siècle, les observations sur la manière de lire *Muḍakkkarāt futūwwa* (vers

1930) par Ḥusni Yūsuf, les considérations théoriques de ‘Uṭmān Ṣabrī dans son introduction à *Rihla fī l-Nīl* (1965) ‘Voyage sur le Nil’, jusqu’à la réflexion de Bayūmī Qandīl sur la ‘āmmiyya dans *Ḥādir al-ṭaqāfa fī miṣr* ‘Le présent de la culture en Égypte’ (1999), etc. De même, les remplacements effectués par Ghalybot pointent vers une normalisation de facto de l’écriture de l’arabe égyptien, selon des principes visant souvent à différencier cette variété de la fuṣḥā. Ces modifications automatiques sont pertinentes pour les pronoms relatifs, par exemple, toutes les formes fuṣḥā pouvant être remplacées par la forme ‘āmmiyya الـ /illi/. Par contre, elles sont parfois impossibles, notamment pour les règles d’écriture de la tā’ marbūṭa, qui dépendent de contextes fluctuants.

Pour répondre, partiellement, à la première question de Florian Coulmas : “What happens when a language is written down, (1) in terms of linguistic description [?]”, il est possible de dire que cela permet une réflexion sur la langue au sens premier du terme, grâce à la visualisation de ce qui était jusqu’ici essentiellement oral. La difficulté à mobiliser une terminologie spécifique pour les phonèmes mis en oeuvre mène cependant souvent à utiliser celle s’appliquant aux graphes, et à confondre “coup de glotte” et hamza, par exemple. Pour répondre à sa deuxième question : “[What happens when a language is written down,] (2) in terms of linguistic evolution ?” on peut estimer que, par son intrusion dans le domaine de l’écrit, la ‘āmmiyya se rapproche de la fuṣḥā qui a été modelée par ce mode d’expression au fil de siècles de pratique. L’emploi des graphes arabes impose aux rédacteurs de l’arabe égyptien d’arbitrer de manière quasi permanente entre écriture étymologique et graphie phonologique. L’objectif est de refléter la prononciation spécifique de la ‘āmmiyya, sans trop s’éloigner de la fuṣḥā à laquelle le lecteur est habitué.

Mais la création d’une édition “dialectale” de Wikipédia est aussi un acte politique, qui a des conséquences sur l’écriture elle-même. Avec l’effacement de hamza sur et sous ‘alif et des deux points de la tā’ marbūṭa et de la yā’ finale, les concepteurs de Wikipédia maṣri ont réussi à donner à l’écriture utilisée dans leur projet une apparence spécifique qui contribue à la différencier de celle de Wikipédia al-‘arabiyya. Cette tendance à simplifier l’écriture n’est pas sans rappeler le style courant de l’écriture manuscrite qui omet régulièrement ces marques pour des raisons de vitesse d’écriture. Cela rappelle les racines populaires de l’arabe égyptien écrit, mis sur le papier majoritairement de manière manuscrite à l’origine. Effacer des éléments de l’écriture renvoie d’une certaine manière aux origines de l’écriture arabe, lorsque les points n’étaient pas utilisés pour discriminer les lettres de même forme. Cela contraste avec les versions modernes du Coran qui non seulement comportent des points mais sont aussi entièrement vocalisées. Ce contraste semble s’inscrire dans l’opposition entre les concepts de Zāhir et Bāṭin mentionnés dans le Coran (57:3) que l’on pourrait traduire par le Manifeste, et l’Invisible. C’est une opposition qui imprègne de multiples aspects de la société arabe, que l’on retrouve dans le croissant de lune repris comme symbole du Croissant rouge par exemple, la partie éclairée de la Lune côtoyant sa partie sombre. Paradoxalement, les concepteurs de Wikipédia maṣri ont ainsi recours à des stratégies traditionnelles de la culture arabe, dont ils cherchent pourtant à se différencier.

Malgré les contradictions inhérentes à un tel projet, cette édition de Wikipédia en arabe égyptien est une expérience précieuse. Tout d’abord, elle a permis une réflexion sur la langue, et la rédaction de règles d’écritures rarement exposées explicitement. Elle a donné lieu à une mobilisation qui contribue à enrichir le corpus écrit de cette variété. On peut regretter l’insuffisance d’application du principe de stigmergie, où la trace laissée dans

l'environnement par l'action initiale stimule une action ultérieure créant une coordination indirecte entre agents⁶⁸. Il reste que cette édition existe toujours, près de dix-sept ans après sa création, et que certains de ses contenus sont uniques sur internet, notamment de nombreux textes de chansons ou de proverbes égyptiens. C'est d'ailleurs là où l'on trouve la plus grande créativité graphique de l'arabe égyptien sur Wikipédia.

L'observation de cette créativité graphique a permis à cette étude de décrire plusieurs phénomènes de mise à l'écrit d'une langue essentiellement parlée. Cela l'inscrit dans un mouvement scientifique lancé par Farnawany (1981), Vial (1983), Badawi et Hinds (1986), Rosenbaum (2004) et Avallone (2015). Un développement souhaitable serait une mise en évidence, non seulement des emplois des graphes, mais aussi des habitudes ou usages pratiqués par ceux qui écrivent l'arabe égyptien, à partir d'un autre type de corpus, tels que des textes publiés. Cela pourrait être une étude diachronique, avec comparaison des pratiques au fil des décennies, ou alors une étude synchronique, portant probablement sur la période actuelle où la publication d'œuvres partiellement ou entièrement en arabe égyptien se fait de manière bien plus décomplexée qu'au siècle précédent. Ceci permettrait d'établir un tableau précis des phénomènes d'écriture d'une langue orale à l'aide d'outils graphiques optimisés pour une variété différente. Cette tension entre orthographe étymologique ou morphologique et graphie phonologique existe déjà en *fuṣḥā*⁶⁹. Elle s'accompagne d'une tension entre les règles morphosyntaxiques de la *fuṣḥā* et les usages différents de la *ʿāmmiyya*, comme par exemple la différence entre les ordres verbe-sujet-objet et sujet-verbe-objet. Ce pourrait être l'esquisse d'une grammaire descriptive de l'arabe égyptien écrit contemporain qui mettrait en évidence des usages implicites qui s'orientent vers une nouvelle norme d'écriture de l'arabe dans sa variété pratiquée en Égypte.

⁶⁸Ce terme a été proposé par le biologiste français Pierre-Paul Grassé en 1959, pour décrire le comportement collaboratif des termites.

⁶⁹Il suffit de rappeler la différence de traitement de *بداية* /bida:ja/ 'début' par rapport à *قراءة* /qira:ʔa/ 'lecture', ou la graphie morphologique de *أنا* ou *في*, dont la voyelle finale est courte à l'oral (avant un article pour ce dernier) et pourtant transcrite par des graphes de voyelles longues.

Bibliographie

- Abdel-Massih, Ernest T., Zaki N. Abdel Malek et El-Said M. Badawi (2009). *A Reference Grammar of Egyptian Arabic*. Washington, DC : Georgetown University Press. url : https://archive.org/details/referencegrammar0000abde_g8y1/mode/2up.
- Aldrich, Matthew (2012). *Egyptian Colloquial Arabic Verbs, Conjugation Tables and Grammar*. Lingualism.
- (2020). *Egyptian Colloquial Arabic Vocabulary*. Arabic Vocabulary. Lingualism.com. url : <https://books.google.fr/books?id=mFfPDwAAQBAJ>.
- Amīn, Muḥammad Šawqī. et Ibrahim Tarzi, éd. (1984). *Majma‘ al-luġa al-‘arabiyya bil-qāhira* - مجمع اللغة العربية بالقاهرة. Le Caire : الهيئة العامة لشئون المطابع الأميرية. url : <https://ebook.univeyes.com/105940/pdf-%D9%85%D8%AC%D9%85%D9%88%D8%B9%D8%A9-%D8%A7%D9%84%D9%82%D8%B1%D8%A7%D8%B1%D8%A7%D8%AA-%D8%A7%D9%84%D8%B9%D9%84%D9%85%D9%8A%D8%A9-%D9%81%D9%8A-%D8%AE%D9%85%D8%B3%D9%8A%D9%86-%D8%B9%D8%A7%D9%85%D8%A7>.
- Ammar, Sam et Joseph Dichy (2008). *Arabe - Les verbes, Bescherelle*. Paris : Hatier.
- Asākir, Ḥalīl (1955). «Ṭarīqa li-kitāba nuṣūṣ il-lahġāt il-‘arabiyya il-ḥadīṭa biḥurūf ‘arabiyya». In : *Mağalla maġma‘ il-luġa il-‘arabiyya*, p. 181-192. url : <https://archive.alsharekh.org/Articles/321/21179/480033>.
- Attia, ‘Adil (oct. 2010). «Yā’ al-wazza ! - يا الوزة !» In : *Al-ḥiwār al-mutamaddin* - الحوار المتمدن 3145. url : <https://www.ahewar.org/debat/show.art.asp?aid=230989>.
- Attia, Mohammed et al. (2015). «Arabic Spelling Error Detection and Correction». In : *Natural Language Engineering* 25. doi : 10.1017/S1351324915000030.
- Avallone, Lucia (2015). «Spelling variants in written Egyptian Arabic, a Study on Literary Texts». In : *Arabic Varieties : Far and Wide - Proceedings of the 11th International Conference of AIDA*, p. 79-86. url : https://www.academia.edu/34816495/Spelling_variants_in_written_Egyptian_Arabic.
- Badawi, El-Said M. (1973). *Mustawayāt al-‘Arabīyah al-mu‘āṣirah fī Miṣr* - مستويات العربية المعاصرة في مصر. Dār al-Ma‘ārif - دار المعارف. url : https://archive.org/details/phonetics_20160228.
- Badawi, El-Said M., Michael Carter et Adrian Gully (2016). *Modern Written Arabic : A Comprehensive Grammar*. New-York : Routledge. url : https://books.google.fr/books?id=VUHVdTk598EC&printsec=frontcover&redir_esc=y#v=onepage&q&f=false.

- Badawi, El-Said M. et Martin Hinds (1986). *A dictionary of Egyptian Arabic : Arabic-English*. Librairie du Liban. url : <https://archive.org/details/ADictionaryOfEgyptianArabic/mode/2up>.
- Baker, M.A. (2021). *How I wrote a million Wikipedia articles*. Independently Published. isbn : 9798716350724. url : <https://dam-oclc.bac-lac.gc.ca/download?id=392610be-1b30-4c5a-94fb-2be4c774dd2c&fileName=How%20I%20wrote%20a%20million%20Wikipedia%20articles.pdf>.
- Behnstedt, P. et M. Woidich (1987). *Die Agyptisch-Arabischen Dialekte I Und II : 1. Einleitung Und Anmerkung Zu Den Karten. 2. Dialektatlas Von Agypten*. Beihefte zum Tübinger Atlas des Vorderen Orients : Geisteswissenschaften. L. Reichert. isbn : 9783882264159.
- Benveniste, Émile. (1966). *Problèmes de linguistique générale*. Problèmes de linguistique générale v. 1. Gallimard. url : <https://books.google.fr/books?id=tqTqngEACAAJ>.
- Björnsson, Jan Arild (nov. 2010). «Egyptian Romanized Arabic : A Study of Selected Features from Communication Among Egyptian Youth on Facebook». Mém. de mast. University of Oslo. url : <https://www.duo.uio.no/handle/10852/24250>.
- Boutros, Wadie (1993). *Ahlan wa sahlān, méthode d'arabe égyptien du Caire*. Le Caire : Éditions Dar el-Nashr Hatier.
- (2000). *Lexique franco-égyptien : le parler du Caire*. Bibliothèque Générale - Institut Français d'Archéologie Orientale. Institut français d'archéologie orientale. isbn : 9782724702712. url : <https://archive.org/details/lexiquefrancoegy0000bout/mode/2up>.
- Catach, Nina (1979). «Le graphème». In : *Pratiques* 25.1. Included in a thematic issue : Orthographe, p. 21-32. doi : 10.3406/prati.1979.1122.
- Coulmas, Florian (2003). *Writing Systems : An Introduction to Their Linguistic Analysis*. Cambridge Textbooks in Linguistics. Cambridge University Press. isbn : 9780521787376. url : <https://books.google.fr/books?id=kmKLxzTnL9IC>.
- Crovitz, Darren et W. Scott Smoot (2009). «Wikipedia : Friend, Not Foe». In : *The English Journal* 98.3, p. 91-97. url : <https://www.jstor.org/stable/40503515>.
- Damardāsh, Ḥ.A.A. (2002). *Waqāyī ' Miṣr al-Qāhira al-Maḥrūsa - وقائع مصر القاهرة المحروسة*. Sous la dir. de M.H. Ibrāhīm et Ṣ.A. Harīdī. Dār Al-Kutub. isbn : 9789771801962. url : <https://books.google.fr/books?id=-bWfAAAAAAAJ>.
- Danyal, M.I. (2015). *Three Shadow Plays by Muhammad Ibn Dānyāl*. Sous la dir. de P. Kahle et D. Hopwood. Gibb Memorial Trust Arabic Studies vol. 32. E.J.W. Gibb Memorial Trust. isbn : 9781909724624. url : <https://books.google.fr/books?id=5sJzzwEACAAJ>.
- Davies, Humphrey (2005). «Dialect Literature». In : *Encyclopedia of Arabic Language and Linguistics*. 1, p. 597-604. url : <https://archive.org/details/the-encyclopedia-of-arabic-language-and-linguistics>.
- Diem, Werner (2006). *Hochsprache und Dialekt im Arabischen : Untersuchungen zur heutigen arabischen Zweisprachigkeit. Abhandlungen für die Kunde des Morgenlandes*. Wiesbaden : Franz Steiner Verlag.

- Doss, Madiha (1996). «Réflexions sur les débuts de l'écriture dialectale en Égypte». In : *Égypte / Monde arabe* 27/28, p. 119-145. doi : <https://doi.org/10.4000/ema.1928>.
- Dowaidar, Ibrahim (oct. 2020). «Political Humor in Ibn Mammātī's Kitāb al-Fāshūsh fi Ahkām Qarāqūsh (The Decisions of Qarāqūsh)». In : *Open Linguistics* 6, p. 482-511. doi : 10.1515/opli-2020-0029.
- Fahmī, 'Abd al-'Azīz (1944). *Al-Ḥurūf al-latīniyya li-kitābat al-'arabiyya* - الحروف اللاتينية لكتابة العربية. Hindawi. url : <https://www.hindawi.org/books/17907319/>.
- Farnawany, Refaat (1981). «Ägyptisch-Arabisch als Geschriebene Sprache, Probleme der Verschriftung einer Umgangssprache». Inaugural Dissertation sur Erlangung des Doktorgrades der Philosophischen Fakultät II. Friedrich Alexander universität.
- Ferguson, Charles A. (1959a). «Diglossia». In : *Word* 15, p. 325-340. url : <https://www.tandfonline.com/doi/pdf/10.1080/00437956.1959.11659702>.
- (1959b). «The Arabic Koine». In : *Language* 35.4, p. 616-630. url : <https://www.jstor.org/stable/410601>.
- Gadalla, Hassan (2000). *Comparative Morphology of Standard and Egyptian Arabic*. 2nd. Munich : Lincom Europa. url : https://www.researchgate.net/publication/369884455_Comparative_Morphology_of_Standard_and_Egyptian_Arabic.
- Girod, Alain et Emad Aziz (2016). *Grammaire de l'arabe égyptien d'aujourd'hui*. Paris : Édition Ellipses.
- Habash, Nizar, Mona Diab et Owen Rambow (mai 2012). «Conventional Orthography for Dialectal Arabic». In : *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. Sous la dir. de Nicoletta Calzolari et al. Istanbul, Turkey : European Language Resources Association (ELRA), p. 711-718. url : http://www.lrec-conf.org/proceedings/lrec2012/pdf/579_Paper.pdf.
- Hanna, Sameh (2019). «When Jesus speaks colloquial Egyptian Arabic : an incarnational understanding of translation». In : *Religion* 49.3, p. 364-387. url : <https://doi.org/10.1080/0048721X.2019.1622841>.
- Haralambous, Yannis (1998). «Simplification of the Arabic Script : Three Different Approaches and their Implementations.» In : *Proceedings of the 7th International Conference on Electronic Publishing, EP'98 Held Jointly with the 4th International Conference on Raster Imaging and Digital Typography, RIDT'98, March 1998, Saint-Malo, France* hal-02101615, p. 138-156. doi : 10.1007/BFb0053268.
- Harrell, Richard S. (1957). *The Phonology of Colloquial Egyptian Arabic*. New York : American Council of Learned Societies.
- Hasan, Heather (2012). *Wikipedia, 3.5 million articles & counting : using and assessing the people's encyclopedia*. New York : Rosen Central. url : <https://archive.org/details/wikipedia35milli0000hasa>.
- Idrīs, Yūsif (2018). *Ġumhūriyyit Farahāt* - "جمهورية فرحات". Hindawi. url : <https://www.hindawi.org/books/48462695/>.
- Imbābi, Faṭḥi (1994). *Marā'ī al-qatl* - مراعي القتل. Al-nahr lilnaṣr wa l-tawzī'. url : <https://www.noor-book.com/%D9%83%D8%AA%D8%A7%D8%A8-%D9%85%D8%B1%D8%A7%D8%B9%D9%8A-%D8%A7%D9%84%D9%82%D8%AA%D9%84-pdf>.
- Jomier, Jacques et Joseph Khouzam (1973). *Manuel d'arabe égyptien*. Paris : Éditions Klincksieck.

- Khalafallah, Abdelghany A. (1969). *A Descriptive Grammar of Saidi Egyptian Colloquial Arabic*. The Hague : Mouton. url : <https://archive.org/details/SaidiArabic>.
- Kouloughli, D.E. (1994). *Grammaire de l'arabe d'aujourd'hui*. Langues pour tous. Pocket. isbn : 9782266039123. url : <https://archive.org/details/software>.
- Lagrange, Frédéric (2013). «Jeux diglossiques et métalinguistiques dans le Qanṭara lladī kafar de Muṣṭafā Muṣarrafa, premier roman en arabe dialectal cairote». In : *Arabica* T. 60, p. 743-761. url : https://www.academia.edu/14144723/Jeux_diglossiques_et_m%C3%A9talinguistiques_dans_le_Qan%E1%B9%ADara_lla%E1%B8%8F%C4%AB_kafar_de_Mu%E1%B9%A3%E1%B9%ADaf%C4%81_Mu%C5%A1arrafa_premier_roman_en_arabe_dialectal_cairote.
- Larcher, Pierre (2001). «Moyen arabe et arabe moyen». In : *Arabica* T. 48, Fasc. 4, Linguistique Arabe : Sociolinguistique et Histoire de la Langue, p. 578-609. url : <https://www.jstor.org/stable/4057670>.
- Louis, Samia (2007). *Kallimni 'Arabi*. Cairo : The American University in Cairo Press. url : <https://archive.org/details/kallimni-arabi-1/mode/2up>.
- Mace, John (1999). *Teach Yourself Arabic Verbs and Essential Grammar*. NTC/Contemporary Pub. url : <https://archive.org/details/arabicverbsessen00john>.
- Miannay, Denis (1972). «Les auditeurs de l'ORTF et la création lexicale». In : *La Banque des mots* 3, p. 3-16.
- Michalski, Marcin (2019). *Written Moroccan Arabic. A Study of Qualitative Variational Heterography*. Poznań : Wydawnictwo Naukowe UAM. url : https://www.academia.edu/45416623/Written_Moroccan_Arabic_A_study_of_qualitative_variational_heterography.
- Mitchell, T. F. (1956). *An Introduction to Egyptian Colloquial Arabic*. Oxford : Oxford University Press. url : https://books.google.fr/books/about/An_Introduction_to_Egyptian_Colloquial_Ar.html?id=gpduwAEACAAJ&redir_esc=y.
- Mūsā, Salāma (juill. 1926). «al-luḡa al-fuṣḥā w al-luḡa al-‘āmmiyya - اللغة الفصحى واللغة العامية». In : *al-Hilāl* - الهلال 10, p. 1073-1077. url : <https://archive.alsharekh.org/Articles/134/13186/259832>.
- Nabulssi-Maselbas, Zuzanna (2023). «Egyptian Linguistic Separatism, A Study in Wikipedia Masri». Varsovie : Université de Varsovie. url : https://www.researchgate.net/publication/378243961_EGYPTIAN_LINGUISTIC_SEPARATISM_A_Study_in_Wikipedia_Masri.
- al-Naqqach, Raga' (1991). «Youssef Idris : Vivre dangereusement». In : *Égypte/Monde arabe* 7, p. 151-155. url : <https://journals.openedition.org/ema/1181>.
- Panović, Ivan (2010). «The Beginnings of Wikipedia Masry». In : *al-Logha Series of Papers in Linguistics* 8, p. 93-127. url : https://www.academia.edu/363979/The_Beginnings_of_Wikipedia_Masry.
- Parkinson, Dilworth B (oct. 1990). «Orthographic variation in modern standard arabic : the case of the Hamza». In : *Perspectives on Arabic Linguistics : Papers from the Annual Symposium on Arabic Linguistics. Volume II : 1988*. Sous la dir. de M. Eid et J. McCarthy. Perspectives on Arabic Linguistics. Salt Lake City, Utah : John Benjamins

- Publishing Company, p. 269-295. isbn : 9789027278289. url : <https://books.google.fr/books?id=uU1AAAAQBAJ>.
- Putten, Marijn van (2017). «The Feminine Ending -at as a Diptote in the Qur'ānic Consonantal Text and Its Implications for Proto-Arabic and Proto-Semitic». In : *Arabica* 64.5-6, p. 695-705. doi : 10.1163/15700585-12341474. url : https://brill.com/view/journals/arab/64/5-6/article-p695_695.xml.
- Qandīl, Bayūmī (1999). *Hādir al-taqāfa fī miṣr - حاضر الثقافة في مصر*. Alexandrie : Dār al-wafā'. url : <https://archive.org/details/BayomiAndil-HadirElThakafa>.
- Reig, Daniel (1984). *'Alf fi 'l wa fi 'l. Kitāb al-taṣrīf. La conjugaison arabe*. Paris : Maisonneuve et Larose.
- Roper, Geoffrey (2014). «Arabic Printing : Printing Culture in the Islamic Context». In : *Encyclopedia of Mediterranean Humanism*. Sous la dir. d'Houari Touati. url : <http://www.encyclopediae-humanisme.com/?Arabic-printing> (visité le 22/12/2023).
- Rosenbaum, Gabriel M. (2000). «Fuṣḥāmmiyya : alternating style in Egyptian prose». In : *Zeitschrift für arabische Linguistik* 38, p. 68-87. url : <https://www.jstor.org/stable/43526795>.
- (2004). «Egyptian Arabic as a written language». In : *Jerusalem Studies in Arabic and Islam*, p. 281-340.
- (2014). «The Growing Influence of English on Egyptian Arabic». In : *Alf lahğa wa-lahğa : Proceedings of the 9th AIDA Conference*. (Neue Beihefte zur Wiener Zeitschrift für die Kunde des Morgenlandes). Vienna and Berlin : LIT Verlag, p. 377-385.
- (2017). «Discussing the use of dialect versus standard Arabic in Lenīn al-Ramlī's Ḥaṣāwi-skits and other texts». In : *Jerusalem Studies in Arabic and Islam* 44, p. 37-66.
- Rubenson, Samuel (juill. 1996). «The Transition from Coptic to Arabic». In : *Égypte/Monde arabe* 27-28. url : <http://journals.openedition.org/ema/1920>.
- Sannū', Ya'qūb (2019). *Mūlyir Miṣr wa ma yuqāsih - مولير مصر وما يقاسيه*. Hindawi. url : <https://www.hindawi.org/books/36308593/>.
- Sartori, Manuel (2014). «La langue des manuscrits grammaticaux arabes médiévaux. Entre fuṣḥā et 'āmmiyya». In : *Romano-Arabica* 14, p. 301-317. url : <https://shs.hal.science/halshs-00961361/document>.
- Al-Sāwī, Muḥammad (2017). *Kitāba al-'arabiyya bil-hurūf al-lātiniyya - كاتبة العربية بالحروف اللاتينية*. Noor Publishing. url : https://www.academia.edu/843265/writing_Arabic_in_the_Latin_letters_%D9%83%D8%AA%D8%A7%D8%A8%D8%A9_%D8%A7%D9%84%D8%B9%D8%B1%D8%A8%D9%8A%D8%A9_%D8%A8%D8%A7%D9%84%D8%AD%D8%B1%D9%88%D9%81_%D8%A7%D9%84%D9%84%D8%A7%D8%AA%D9%8A%D9%86%D9%8A%D8%A9.
- Shirbīnī, Y.M. et H.T. Davies (2005). *Yūsuf Al-Shirbīnī's Kitāb Hazz Al-quḥūf Bi-Sharḥ Qaṣīd Abī Shādūf : A lexicon of 17th-century Egyptian Arabic*. Orientalia Lovaniensia analecta. Peeters. isbn : 9789042915275. url : <https://books.google.fr/books?id=TnQozQEACAAJ>.
- Spitta, Wilhelm (1880). *Grammatik des Arabischen Vulgärdialectes von Aegypten*. Leipzig : J. C. Hinrichs. url : <https://archive.org/details/grammatikdesara00spitgoog/mode/2up>.

- Stephens, Virginia et Salib Maurice (2004). *A Pocket Dictionary of the Spoken Arabic of Cairo*. Cairo : The American University in Cairo Press.
- Taymūr, Maḥmūd (1949). *Makhba' raqam 13* - [مخبأ رقم ١٣]. Le Caire : [مطبعة عطايا]. url : <https://foulabook.com/ar/book/%D8%A7%D9%84%D9%85%D8%AE%D8%A8%D8%A3-%D8%B1%D9%82%D9%85-13-pdf>.
- Al-Toma, Salih J. (1961). «The Arabic Writing System and Proposals for Its Reform». In : *Middle East Journal* 15.4, p. 403-415. url : <https://www.jstor.org/stable/4323405> (visité le 22/12/2023).
- Al-Tonsi, Abbass, Laila Al-Sawi et Suzanne Massoud (2010). *Kalaam Gamiil, An Intensive Course in Egyptian Colloquial Arabic*. T. 1. Le Caire – New York : The American University in Cairo Press. url : https://books.google.fr/books/about/Kalaam_Gamiil.html?id=SCOBwAEACAAJ&redir_esc=y.
- Vial, Charles (1983). *L'égyptien tel qu'on l'écrit : glossaire établi d'après un choix d'œuvres littéraires égyptiennes contemporaines*. L'égyptien tel qu'on l'écrit : glossaire établi d'après un choix d'œuvres littéraires égyptiennes contemporaines v. 18; v. 1983. Institut français d'archéologie orientale du Caire. url : <https://books.google.fr/books?id=MC9BRjYmW7AC>.
- Vollers, Karl (1890). *Lehrbuch der aegypto-arabischen Umgangssprache*. auto-publié par l'auteur.
- (1895). *The Modern Egyptian Dialect of Arabic : A Grammar With Exercises, Reading Lessons and Glossaries*. Cambridge University Press.
- Watson, Janet C. E. (2002). *The Phonology and Morphology of Arabic*. Oxford University Press.
- Wehr, H. et J.M. Cowan (1979). *A Dictionary of Modern Written Arabic*. Harrassowitz. isbn : 9783447020022. url : <https://books.google.fr/books?id=WTak55pG-IC>.
- Willmore, J. S. (1901). *The Spoken Arabic of Egypt*. London : Dalcassian Publishing Company.
- Woidich, M. (2006). *Das Kairenisch-Arabisches : eine Grammatik*. Porta linguarum orientaliū. Harrassowitz.
- Woidich, Manfred (1994). «Cairo Arabic and the Egyptian Dialects». In : *Language and Literature - LANG LIT.*, p. 493-507.
- (juill. 2008). «Rural Dialect of Egyptian Arabic : An Overview». In : *Égypte/Monde arabe*, p. 27-28. url : <http://journals.openedition.org/ema/1952>.
- (2010). «Von der wörtlichen Rede zur Sachprosa : Zur Entwicklung der Ägyptisch-Arabischen Dialektliteratur». In : *Dialektliteratur heute – regional und international. Forschungskolloquium am Interdisziplinären Zentrum für Dialektforschung an der Friedrich-Alexander-Universität Erlangen-Nürnberg* 2009, p. 63-94.
- Wright, William (1967). *A grammar of the Arabic language*. T. I. Cambridge : Cambridge University Press.
- Youssef, Islam (2013). «Place Assimilation in Arabic : Contrasts, Features, and Constraints». A dissertation for the degree of Philosophiae Doctor. University of Tromsø.

- Youssef, Islam (jan. 2014). «Emphasis spread in Cairene Arabic : A reassessment». In : *Alf lahġa wa-lahġa : Proceedings of the 9th AIDA Conference*, p. 455-464. isbn : 978-3-643-90334-1.
- (2021). «Contrastive Feature Typologies of Arabic Consonant Reflexes». In : *Languages* 6, p. 141. doi : 10.3390/languages6030141.
- Zack, Liesbeth (2001). «The use of colloquial arabic in prose literature : Laban il'āsfūr by Yūsuf al-Qa'id». In : *Quaderni di Studi Arabi* 19, p. 193-219.
- (2014). «Key to Mass Literacy or Professor's Hobby ? Fiske's Project to Write Egyptian Arabic with the Latin Alphabet». In : *Al-Arabiyya* 47.2014, p. 1-19. url : <https://www.jstor.org/stable/24635370>.
- Al-Zoubi, Samer (jan. 2019). «The Speech Sounds of Arabic Language and their Effect on Learning English Pronunciation : A Contrastive Analysis». In : *International Journal of Humanities and Social Science* 9. doi : 10.30845/ijhss.v9n1p2.

Pages internet

- ISO (2024). *ISO 639-3* | — iso639-3.sil.org. [Consulté le 20/03/2024]. url : <https://iso639-3.sil.org/>.
- Semrush (2025). *Most Visited Websites in Worldwide 2025* | *Open .Trends* — [semrush.com](https://www.semrush.com/trending-websites/global/all). [Consulté le 14/04/2025]. url : <https://www.semrush.com/trending-websites/global/all>.
- Wikidata (2023). *Wikidata* — [m.wikidata.org](https://m.wikidata.org/wiki/Wikidata:Main_Page?uselang=fr). [Consulté le 14/04/2025]. url : https://m.wikidata.org/wiki/Wikidata:Main_Page?uselang=fr.
- Wikimedia (2018a). *Requests for new languages/Wikipedia Egyptian Arabic - Meta* — [meta.wikimedia.org](https://meta.wikimedia.org/wiki/Requests_for_new_languages/Wikipedia_Egyptian_Arabic). [Consulté le 5/03/2024]. url : https://meta.wikimedia.org/wiki/Requests_for_new_languages/Wikipedia_Egyptian_Arabic.
- (2018b). *Wikimedia Traffic Analysis Report - Wikipedia Page Views Per Country - Breakdown* — [stats.wikimedia.org](https://stats.wikimedia.org/wikimedia/squids/SquidReportPageViewsPerCountryBreakdown.htm). [Consulté le 12/03/2024]. url : <https://stats.wikimedia.org/wikimedia/squids/SquidReportPageViewsPerCountryBreakdown.htm>.
- (2024). *Wikimedia Traffic Analysis Report - Page Views Per Wikipedia Language - Breakdown* — [stats.wikimedia.org](https://stats.wikimedia.org/wikimedia/squids/SquidReportPageViewsPerLanguageBreakdown.htm). [Consulté le 12/03/2024]. url : <https://stats.wikimedia.org/wikimedia/squids/SquidReportPageViewsPerLanguageBreakdown.htm>.
- (2025a). *Language proposal policy*. [Consulté le 14/04/2025]. url : https://meta.wikimedia.org/wiki/Language_proposal_policy.
- (2025b). *List of Wikipedias*. [Consulté le 14/04/2025]. url : https://meta.wikimedia.org/wiki/List_of_Wikipedias.
- (2025c). *List of Wikipedias by speakers per article - Meta* — [meta.wikimedia.org](https://meta.wikimedia.org/wiki/List_of_Wikipedias_by_speakers_per_article). [Consulté le 14/04/2025]. url : https://meta.wikimedia.org/wiki/List_of_Wikipedias_by_speakers_per_article.
- (2025d). *Règlement des propositions de langues*. [Consulté le 14/04/2025]. url : https://meta.wikimedia.org/wiki/Language_proposal_policy/fr.

- Wikimedia (2025e). *Wikipedia article depth - Meta* — *meta.wikimedia.org*. [Consulté le 14/04/2025]. url : https://meta.wikimedia.org/wiki/Wikipedia_article_depth.
- (2025f). *Wikistats - Statistics For Wikimedia Projects* — *stats.wikimedia.org*. [Consulté le 14/04/2025]. url : <https://stats.wikimedia.org/#/arz.wikipedia.org>.
- Wikipédia (2020a). نقاش ويكيبيديا:طريقة الكتابة/ارشيف 1 - ويكيبيديا — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : https://arz.wikipedia.org/wiki/%D9%86%D9%82%D8%A7%D8%B4_%D9%88%D9%8A%D9%83%D9%8A%D8%A8%D9%8A%D8%AF%D9%8A%D8%A7:%D8%B7%D8%B1%D9%8A%D9%82%D8%A9_%D8%A7%D9%84%D9%83%D8%AA%D8%A7%D8%A8%D9%87/%D8%A7%D8%B1%D8%B4%D9%8A%D9%81_1.
- (2020b). نقاش ويكيبيديا:طريقة الكتابة/ارشيف 2 - ويكيبيديا — *arz.wikipedia.org*. [Consulté le 11/07/2024]. url : https://arz.wikipedia.org/wiki/%D9%86%D9%82%D8%A7%D8%B4_%D9%88%D9%8A%D9%83%D9%8A%D8%A8%D9%8A%D8%AF%D9%8A%D8%A7:%D8%B7%D8%B1%D9%8A%D9%82%D8%A9_%D8%A7%D9%84%D9%83%D8%AA%D8%A7%D8%A8%D9%87/%D8%A7%D8%B1%D8%B4%D9%8A%D9%81_2.
- (2020c). نقاش ويكيبيديا:طريقة الكتابة/ارشيف 3 - ويكيبيديا — *arz.wikipedia.org*. [Consulté le 11/07/2024]. url : https://arz.wikipedia.org/wiki/%D9%86%D9%82%D8%A7%D8%B4_%D9%88%D9%8A%D9%83%D9%8A%D8%A8%D9%8A%D8%AF%D9%8A%D8%A7:%D8%B7%D8%B1%D9%8A%D9%82%D8%A9_%D8%A7%D9%84%D9%83%D8%AA%D8%A7%D8%A8%D9%87/%D8%A7%D8%B1%D8%B4%D9%8A%D9%81_3.
- (2020d). *HitomiAkane/arzAutocorrect.js* — *مستخدم* — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : <https://arz.wikipedia.org/wiki/%D9%85%D8%B3%D8%AA%D8%AE%D8%AF%D9%85:HitomiAkane/arzAutocorrect.js>.
- (2021). *MatthewS:مستخدم* — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : https://arz.wikipedia.org/wiki/%D9%85%D8%B3%D8%AA%D8%AE%D8%AF%D9%85:MatthewS./%D8%A7%D9%84%D9%81-%D8%A8%D9%87_%D9%85%D8%B5%D8%B1%D9%89_%D9%84%D8%A7%D8%AA%D9%8A%D9%86%D9%89.
- (2024a). اللغة المصريه الحديثه - ويكيبيديا — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : https://arz.wikipedia.org/wiki/%D8%A7%D9%84%D9%84%D8%BA%D9%87_%D8%A7%D9%84%D9%85%D8%B5%D8%B1%D9%8A%D9%87_%D8%A7%D9%84%D8%AD%D8%AF%D9%8A%D8%AB%D9%87.
- (2024b). *GhalyBot* — *مستخدم* — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : <https://arz.wikipedia.org/wiki/%D9%85%D8%B3%D8%AA%D8%AE%D8%AF%D9%85:GhalyBot>.
- (2024c). *Main Page* — *Wikipedia, The Free Encyclopedia*. [Consulté le 5/03/2024]. url : <http://en.wikipedia.org/w/index.php?title=Main%5C%20Page&oldid=1225315602>.
- (2025a). داربجة - ويكيبيديا — *ary.wikipedia.org*. [Consulté le 08/03/2025]. url : <https://ary.wikipedia.org/wiki/%D8%AF%D8%A7%D8%B1%D9%8A%D8%AC%D8%A9>.
- (2025b). مقالات باللاتيني - ويكيبيديا — *arz.wikipedia.org*. [Consulté le 14/04/2025]. url : https://arz.wikipedia.org/wiki/%D9%85%D9%82%D8%A7%D9%84%D8%A7%D8%AA_%D8%A8%D8%A7%D9%84%D9%84%D8%A7%D8%AA%D9%8A%D9%86%D9%89.

- Wikipédia (2025c). *ويكيبيديا:طريقة الكتابة - ويكيبيديا*. — [arz.wikipedia.org](https://arz.wikipedia.org/wiki/%D9%88%D9%8A%D9%83%D9%8A%D8%A8%D9%8A%D8%AF%D9%8A%D8%A7:%D8%B7%D8%B1%D9%8A%D9%82%D8%A9_%D8%A7%D9%84%D9%83%D8%AA%D8%A7%D8%A8%D9%87). [Consulté le 14/04/2025]. url : https://arz.wikipedia.org/wiki/%D9%88%D9%8A%D9%83%D9%8A%D8%A8%D9%8A%D8%AF%D9%8A%D8%A7:%D8%B7%D8%B1%D9%8A%D9%82%D8%A9_%D8%A7%D9%84%D9%83%D8%AA%D8%A7%D8%A8%D9%87.
- (2025d). *Egyptian Arabic Wikipedia*. [Consulté le 14/04/2025]. url : https://en.wikipedia.org/wiki/Egyptian_Arabic_Wikipedia.
- (2025e). *Ghaly/autowikiGhalyBot[:مستخدم]*. — [arz.wikipedia.org](https://arz.wikipedia.org/wiki/%D9%85%D8%B3%D8%AA%D8%AE%D8%AF%D9%85:Ghaly/autowikiGhalyBot). [Consulté le 14/04/2025]. url : <https://arz.wikipedia.org/wiki/%D9%85%D8%B3%D8%AA%D8%AE%D8%AF%D9%85:Ghaly/autowikiGhalyBot>.
- (2025f). *GhalyBot[:مساهمات اليوزر ل]*. — [arz.wikipedia.org](https://arz.wikipedia.org/wiki/%D8%AE%D8%A7%D8%B5:%D9%85%D8%B3%D8%A7%D9%87%D9%85%D8%A7%D8%AA/GhalyBot). [Consulté le 14/04/2025]. url : <https://arz.wikipedia.org/wiki/%D8%AE%D8%A7%D8%B5:%D9%85%D8%B3%D8%A7%D9%87%D9%85%D8%A7%D8%AA/GhalyBot>.
- (2025g). *Wikipédia :Bot*. [Consulté le 14/04/2025]. url : <https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Bot>.
- WorldBank (2024). *Health Nutrition and Population Statistics | DataBank*. — data-bank.worldbank.org. [Consulté le 12/03/2024]. url : <https://data-bank.worldbank.org/reports.aspx?source=health-nutrition-and-population-statistics>.