

Catherine Miller, Alexandrine Barontini, Marie-Aimée Germanos, Jairo Guerrero and Christophe Pereira (dir.)

Studies on Arabic Dialectology and Sociolinguistics
Proceedings of the 12th International Conference of AIDA held in
Marseille from May 30th to June 2nd 2017

Institut de recherches et d'études sur les mondes arabes et musulmans

Communication écrite sur Internet et par sms en *arabizi* : analyse de corpus relevant des glosses dialectales libanaise et marocaine

Joseph Dichy, Fatima-Zahra Driss, Hassiba Lagraf et Djamel Mostefa

DOI : 10.4000/books.iremam.4811

Éditeur : Institut de recherches et d'études sur les mondes arabes et musulmans

Lieu d'édition : Aix-en-Provence

Année d'édition : 2019

Date de mise en ligne : 24 janvier 2019

Collection : Livres de l'IREMAM

ISBN électronique : 9791036533891



<http://books.openedition.org>

Référence électronique

DICHY, Joseph ; et al. *Communication écrite sur Internet et par sms en arabizi : analyse de corpus relevant des glosses dialectales libanaise et marocaine* In : *Studies on Arabic Dialectology and Sociolinguistics : Proceedings of the 12th International Conference of AIDA held in Marseille from May 30th to June 2nd 2017* [en ligne]. Aix-en-Provence : Institut de recherches et d'études sur les mondes arabes et musulmans, 2019 (généré le 12 janvier 2021). Disponible sur Internet : <<http://books.openedition.org/iremam/4811>>. ISBN : 9791036533891. DOI : <https://doi.org/10.4000/books.iremam.4811>.

Ce document a été généré automatiquement le 12 janvier 2021.

Communication écrite sur Internet et par sms en *arabizi* : analyse de corpus relevant des glosses dialectales libanaise et marocaine

Joseph Dichy, Fatima-Zahra Driss, Hassiba Lagraf et Djamel Mostefa

1. Introduction

- 1 La présente contribution porte sur un aspect assez récent de la dialectologie et de la communication en arabe : l'usage de caractères latins pour transmettre des messages, soit par SMS, soit sur Internet. Ce phénomène a été appelé *arabizi*, mot-valise construit à partir de « Arab » ou « Arabic » et « easy ».
- 2 Nos analyses s'inscrivent dans le projet de recherche PARADE (« Plateforme d'Analyse et de Recherche en Arabe et ses DialectEs »), en cours depuis la fin de 2016 (partenaires: SYSTRAN SA [www.systran.com]; Laboratoire ICAR, UMR 5191, CNRS/Université Lyon 2, ENS-Lyon [<http://icar.univ-lyon2.fr>]; ARADIC Editions [www.aradic.fr]). Ce projet porte au premier chef sur l'*arabizi*.
- 3 Nous chercherons à décrire ce dernier à partir de corpus provenant d'Internet et principalement de Twitter. Nous associerons aux informations générales du début des analyses précises d'extraits de corpus, afin de donner une idée des phénomènes en jeu.

1.1. L'explosion d'internet dans les pays arabes

- 4 Globalement, tous les pays arabes ont enregistré une forte hausse du nombre d'utilisateurs d'Internet au cours des 10 dernières années (Figures 1 et 2). La particularité de ces utilisateurs est qu'ils sont très ancrés dans les nouvelles formes de communication telles que les réseaux sociaux, où des pseudonymes leur permettent de s'exprimer librement et dans la forme linguistique de leur choix.

- 5 Le nombre total d'utilisateurs de Facebook dans le monde arabe fin mai 2013 s'élevait à 55 millions, avec une progression annuelle de 22% par rapport à mai 2012 (10 millions de nouveaux utilisateurs en un an). Twitter est également en pleine expansion dans cette aire : l'Arabie Saoudite est le premier pays au monde, en pourcentage, d'utilisateurs de Twitter (peerreach.com, 2013).

Figure 1 : Taux de pénétration d'Internet en pourcentage de la population totale (source ICT Statistics Database)

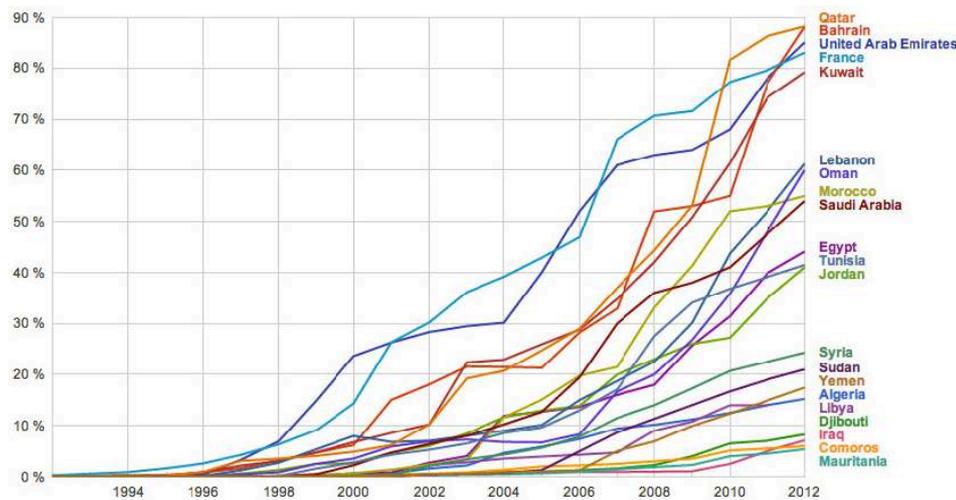
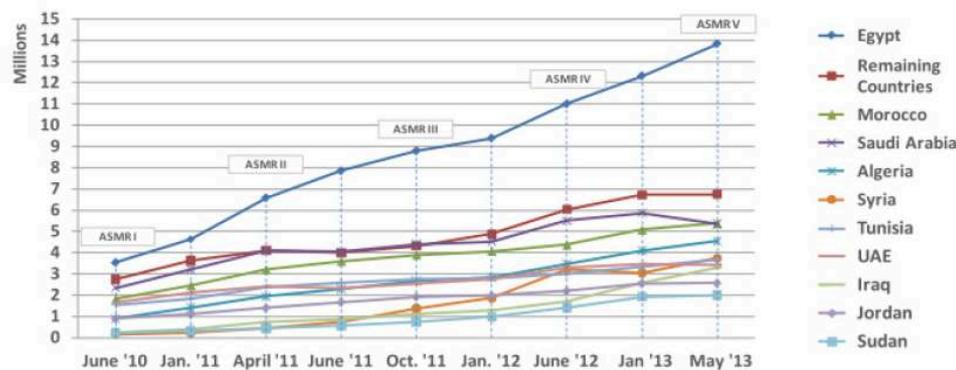


Figure 2 : Nombre d'utilisateurs de Facebook dans le monde arabe (source Arab Social Media Report V)



1.2. Internet dans le monde arabe – formes et variétés linguistiques

- 6 Les utilisateurs d'Internet et des réseaux sociaux ont recours à diverses formes (ou *glosses*) de l'arabe et s'expriment souvent dans des dialectes translittérés. Cela rend l'analyse – et plus encore, l'analyse automatique – de ces données plus difficile. En effet, en fonction de l'origine des auteurs, de leur dialecte ou des claviers utilisés par eux, une forte hétérogénéité apparaît dans les textes transcrits en caractères arabes ou latins, et combinant l'utilisation de variantes dialectales et des mélanges de mots en français ou anglais et en arabe. Les variétés ou glosses de l'arabe ainsi réalisé diffèrent d'un pays à

l'autre et sont influencées par le français au Maghreb et par l'anglais au Machrek (une partie du Liban mise à part), ainsi naturellement que par les dialectes locaux. L'arabizi est en outre utilisé par les enfants de la diaspora arabe qui ne maîtrisent pas l'écriture arabe, mais parlent – peu ou prou – le dialecte de leurs parents, pour communiquer, par exemple, avec leurs relations habitant leur pays d'origine (famille, amis). Un domaine d'analyse relativement nouveau s'offre donc à la dialectologie et à ses descriptions.

- 7 Plusieurs approches existent pour la notion d'arabizi (Denis, Fréard, Détiennne, Baker, & Barcellini, 2012), (Aboelezz, 2009), (Bahrainwala, 2011), (Bianchi, 2012), (Elmahdy, Gruhn, & Minker, 2012), (Yaghan, 2008), (Gonzales-Quijano, 2012 a & b). Sa naissance remonte aux premiers claviers d'ordinateurs et de téléphones portables, qui ne permettaient pas la saisie des caractères arabes. La première apparition de cette forme de notation était dans les SMS, le chat et les courriels informels (Elmahdy et al, 2012).
- 8 Mona Randa, Nariman Elshamly & Nady (2011) ont réalisé une étude sur les utilisateurs de Facebook, montrant que 82% des sondés utilisent l'arabizi, 40% déclarent l'utiliser la plupart du temps et 22% l'utilisent systématiquement. Aboelezz (2009) examine un corpus de 253 courriels provenant de plus de 70 personnes de 16 à 25 ans en Egypte, et montre que 52,4% des messages contenaient de l'arabizi et que ce dernier représentait au moins 60% des messages concernés. Al-Khatib & Sabbah (2008) ont collecté un corpus de SMS provenant d'étudiants jordaniens âgés de 17 à 25 ans dans lequel 61% des messages étaient en arabizi.

2. Le travail sur corpus marocains, syro-libanais et égyptiens

- 9 La présente contribution vise à affiner la définition de l'arabizi à partir d'analyses de corpus. Le projet PARADE porte sur des corpus collectés essentiellement sur Twitter, dans les trois zones dialectales marocaine, syro-libanaise et égyptienne (que nous n'étudions pas dans cette communication).
- 10 Voici de premiers résultats d'observation :
 - On constate, comme on pouvait s'y attendre, des échanges faisant intervenir plusieurs langues et plusieurs types de caractères ou de transcription, c.-à-d., de l'arabe en caractères latins et arabes, mêlé à de l'anglais, du français, de l'espagnol ou de l'italien.
 - On trouve donc aussi de l'arabizi alternant avec des échanges en caractères arabes.
 - Dans ces échanges, l'arabe littéraire moderne peut alterner avec du dialecte.
- 11 En termes linguistiques plus précis, ce qui précède peut être schématisé comme suit :
 - 1) Au sein de la langue arabe considérée dans la multiplicité de ses formes, c.-à-d., au sein de la *polyglossie* de l'arabe, (Dichy 2010), on rencontre dans les échanges en arabizi :
 - deux types d'écriture (caractères arabes ou latins) ;
 - différentes glosses (dialectes, arabes moyens, arabe littéraire moderne).
 - 2) À cela s'ajoutent des phénomènes d'*alternance codique* (en anglais, *code-switching*) entre l'arabe et l'anglais ou le français, voire avec de l'espagnol ou de l'italien.

2.1. Définition « basique » de l'arabizi

- 12 À partir de nos observations, nous essayerons de donner une définition linguistique opératoire de l'arabizi. Citons en première approche la description d'Yves Gonzalez-Quijano, à qui l'on doit par ailleurs de précieuses observations sociologiques :
- « Il s'agit d'une forme écrite de l'arabe née du développement des appareils électroniques : au début, les ordinateurs et autres téléphones portables n'avaient pas de clavier arabe. Les utilisateurs ont donc transposé l'alphabet arabe en alphabet latin. Et quand ils ont buté contre des sons sans équivalent, ils ont introduit des chiffres. Le 3 pour la gutturale 'ayn par exemple, comme dans *3rabya* « arabe » (Gonzalez-Quijano 2012b).
- 13 On remplace ainsi les sons correspondant par exemple aux lettres *ḥā'* ou 'ayn respectivement par 3 et par 7, comme dans *mar7aba*, « bonjour ou bienvenue » et *3ayni* (Liban)/*3eni* (Syrie), « mon œil ». Pour le *ḥā'*, on utilise soit le chiffre 5, soit les deux lettres « kh », pour le *šīn*, on a soit « sh » (régions où l'anglais est en usage), soit « ch » (régions où la langue étrangère est le français). On observe en outre des notations relâchées, intuitives, et donc souvent approximatives.
- 14 La diversité des usages est une caractéristique ancienne de la polyglossie de l'arabe, qui n'en a pas affecté à ce jour l'unité plurielle (Dichy 2010 et 2017 – mais voir aussi les travaux de J. Fück, J. Blau ou J. Lentin sur ce que nous appelons les *arabes moyens*, au pluriel). L'« arabe moyen » étant un mixte d'arabe littéraire moderne et d'un dialecte, il s'ensuit qu'il y a autant d'arabes moyens – au pluriel – que de dialectes.

2.2. L'arabizi se réduit-il à un mode de transcription ?

- 15 Une toute première définition réduit donc l'arabizi à un mode de transcription. Ce dernier utilise des équivalences entre lettres latines pour les sons ou phonèmes communs aux deux langues (arabe algérien et français par exemple) et des chiffres ou couples de lettres pour des phonèmes présents dans le dialecte arabe mais absents du français ou de l'anglais :
- Lettres arabes avec équivalent en caractères latins, exemples : d = dāl, m = mīm.
 - Lettres arabes sans équivalent en alphabet latin et remplacées par des chiffres dont la graphie est en analogie partielle avec la forme de la lettre arabe, exemples : etc. (La ressemblance, pour les exemples précédents est en ح = 7 ; خ = 5 ; ع = 3 « image de miroir ».)
 - Utilisation de digraphes pour certains phonèmes, exemple : dh = ḍāḍ, etc.
- 16 Enfin, les règles de notation ne sont pas univoques : certaines lettres présentent des confusions de transcription, ex: k = *qāf* ou *kāf*. De notables confusions proviennent du caractère peu normé de ces notations. Voici une phrase en arabizi qui mélange lettres latines, chiffres et utilisation de deux lettres pour représenter certains sons : *allah yr7mek ya khali*, « Paix à ton âme, (ô) mon oncle » (dialecte algérien).
- 17 Cette description formelle de base est essentielle. On verra toutefois que l'on ne doit pas s'arrêter à ce niveau d'analyse.

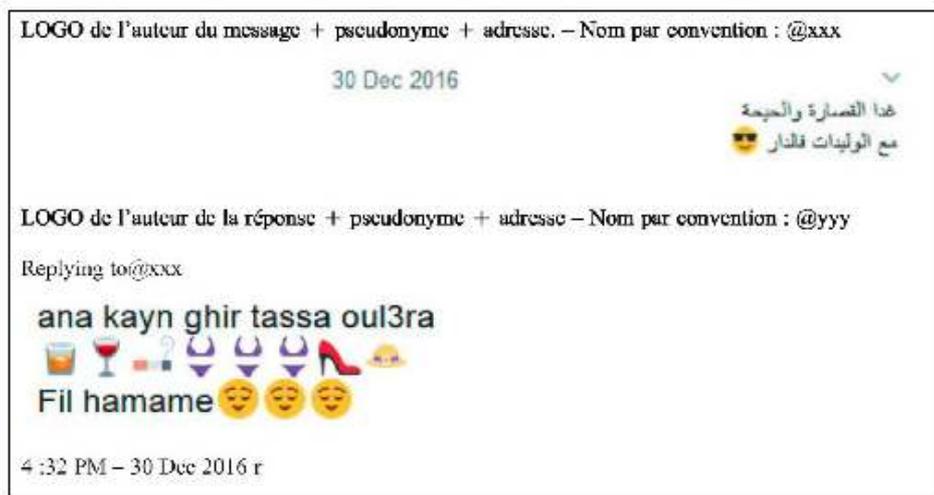
3. Analyse de corpus

- 18 Considérons ci-dessous quatre exemples d'échanges en arabizi sur Twitter : deux extraits marocains et deux extraits libanais. Ces extraits ont été modifiés de manière à

en assurer l'anonymat (suppression des logos, pseudonymes et adresses). Chaque exemple permet d'illustrer un aspect de la communication en arabizi ou des formes qu'elle prend.

3.1. Exemple marocain n°1 : présence de dialecte marocain en graphie arabe et en arabizi, associé à une combinaison d'émoticônes

Figure 3 : Reproduction anonymée du premier extrait



19 Observations sur les données de l'échange

1^{er} interlocuteur, @xxx

Traduction : « Demain on va s'éclater et bavarder [en toute convivialité] avec les enfants à la maison » – « To-morrow, it'll be a great time and cool exchanges with the children at home »

Commentaires : Le texte est en dialectal marocain et en caractères arabes, suivi d'une seule émoticône suggérant l'idée de loisirs et d'un bon moment.

20 2nd interlocuteur, @yyy

Traduction : « Pour moi rien que des verres [d'alcool] et la nudité dans le hammam » – « Me, nothing but drinking glasses [of alcoholic drinks] and nudes in the hammam ». (Le mot noté « tassa » en arabizi, en transcription *ṭāṣa*, désigne la petite bassine utilisée pour s'asperger d'eau chaude ou froide dans le hammam, dans un grand nombre de dialectes arabes (notamment au Proche-Orient). C'est très probablement un emprunt à l'italien. En arabe marocain, ce mot désigne bien un verre à boire, ici de l'alcool, comme l'indique en outre l'une des émoticônes du message).

Commentaires :

(1) Le texte est en arabizi, avec beaucoup d'émoticônes humoristiques, qui expriment une forme d'auto-dérision. Le twitteur @yyy manifeste ainsi que son ami(e) @xxx est en réalité plus heureux(se) que lui, qui est pourtant plus libre, étant apparemment sans enfants. Nous présumons qu'il s'agit d'un homme qui exprime ici, de manière en partie ironique, un fantasme. On peut en outre se demander si la cigarette allumée (3^e émoticône) n'est pas une cigarette de hachich.

(2) Quoiqu'il en soit de nos interprétations, le rôle des émoticônes comme complément sémantique est considérable, et remet partiellement en cause la notion de corpus en arabizi ou en caractères arabes sur réseaux sociaux ou par SMS comme étant uniquement textuel.

(3) L'extrait comporte en outre une combinaison d'émoticônes, incluant des itérations (émoticônes répétées). Rappelons que les émoticônes répondent en partie à la fonction expressive du langage. Mais ils constituent aussi un moyen d'explicitier le sens dans l'interaction (par ex., le smiley pour dire « je plaisante »). Ces itérations expriment :

- soit une insistance de nature exclamative, comme dans merciiaiiiiiii en français, l'itération du symbole correspondant à une façon de mimer par écrit l'intonation ;
- soit l'expression de la quantité par itération (comme dans « des voitures, des voitures, des voitures et des voitures : il n'y a que ça dans cette ville », ou encore : « il est très très très très grand »). Ici, l'image de femmes en sous-vêtements ou en maillot de bain et celle du visage exprimant contentement et satisfaction, sont répétées trois fois.

- 21 Ces deux interprétations peuvent bien entendu s'additionner. Une étude des émoticônes et du langage affectif dans notre corpus serait à faire (voir plus loin, § 3.3), avec des comparaisons à d'autres langues.

3.2. Exemple marocain n°2 : recours à une image ; usage d'arabe moyen marocain en caractères arabes, présence d'implicites conversationnels

Figure 4 : Reproduction anonymée du deuxième extrait



22 Observations sur l'échange

Les interlocuteurs @yyy et @xxx sont les mêmes que dans l'extrait précédent.

1^{er} interlocuteur, @xxx

Commentaire : Le message est entièrement composé d'une image qui présente une bouteille de bière de marque « Flag » posée sur une table dans un café, à côté de deux assiettes. (Flag est une bière marocaine de luxe commercialisée dans plusieurs pays d'Afrique de l'Ouest.)

23 2nd interlocuteur, @yyy

Traduction : « A ta santé et en l'honneur de l'authenticité (*'aṣāla*) et de la modernité (*mu'āṣara*) » – « Tchîn-tchîn (to your good health) and in honour of authenticity and modernity »

Commentaire : Le texte est en caractères arabes. La référence à l'authenticité et à la modernité relève d'un discours « intellectuel » : il y a une forte connotation de ces termes dans la culture arabe contemporaine (problématique de la modernité, du renoncement, ou non, à ses traditions, etc.). Mais on peut se demander également si l'interlocuteur ne fait pas allusion au parti marocain de l'« Authenticité et la Modernité », ce qui donnerait à la réponse un contenu politique, voire politico-religieux, la consommation d'alcool étant prohibée en islam. Le sens sous-entend donc : soit (a), « Quelle chance tu as d'être dans ce bel endroit », où il y a sur la table à côté de la bière un tagine marocain d'accompagnement des consommations ; soit (b) une allusion politico-religieuse ironique, la bière étant elle-même, comme le tagine qui l'accompagne « authentiquement » marocaine, tout en apparaissant comme signe de « modernité », cette dernière étant ici transgressive par rapport à la religion dominante. Les deux interprétations sont probablement vraies en même temps.

24 Réponse du 1^{er} interlocuteur, @xxx

Traduction : « C'est classe, mon frère, classe », « great/splendid, brother, gerat/splendid »

Commentaire : La réponse est en arabizi, et suivie d'une émoticône (bouche féminine), pour dire, nous semble-t-il : « c'est bon, c'est raffiné comme nourriture, boisson et endroit ».

25 Notons que la réponse se fait par rapport à un sens sous-entendu (acte de langage implicite), de ce que dit le 2^e interlocuteur, ce qui est une caractéristique de l'interaction. Ces implicites comportent comme on l'a vu, beaucoup de contenus, dont certains ne sont pas directement accessibles à partir du texte.

26 Si dans l'exemple précédent les émoticônes jouaient un rôle central, on voit ici que dans le premier message, ce rôle est tenu par une image. On voit également que le 2nd locuteur, qui écrit en caractères arabes, faisait usage du parler marocain dans le premier extrait, et d'un arabe moyen dans le second.

3.3. Exemples n° 3 et 4, en arabe libanais

27 Les extraits ci-dessous font apparaître un usage exclusif de l'arabizi, ainsi que des alternances codiques avec l'anglais dans le premier exemple, mais pas du tout dans le second. Le premier extrait comporte une image, le second un logo objet de commentaires acerbes.

28 On observe dans les deux cas une utilisation expressive ou affective de la dimension des caractères (comme aussi pour @xxx dans les extraits marocains). Les grands caractères sont accompagnés d'émoticônes dans l'exemple n°3, mais pas dans l'extrait n°4. On a donc, pour l'expressivité, le schéma de possibilités suivant :

Figure 5 : Fonction expressive des émoticônes en relation avec la dimension des caractères : (schéma des possibilités)

Dimension des caractères et des émoticônes	Emoticône(s)		
	standard	absente	unique
agrandie	absente	unique	itération de l'émoticône

3.4. Exemple n°3, arabe libanais

Figure 6 : Reproduction partiellement anonymée du troisième extrait



29 Observations sur cet extrait

1^{re} interlocutrice, @mmm

Traduction : Mot-à-mot, « La protection de Dieu sur la beauté de ses yeux et de son sourire ».

Commentaire :

(1) Nous n'avons pas anonymisé ce qui concerne l'acteur Joe Sader, car il s'agit de sites « officiels ». Ceux-ci sont très élaborés ; ils fonctionnent par abonnement gratuit et annoncent selon le cas, plusieurs centaines à plusieurs milliers d'abonnés – voir <https://www.instagram.com/joesaderofficial/?hl=fr>.

(2) La première expression est en arabizi. Il s'agit de conjurer le mauvais œil (cette superstition pèse toujours sur les compliments, concernant notamment les yeux). L'usage de l'arabe s'impose donc, l'alternance codique avec l'anglais intervenant juste

après. Suivent deux émoticônes exprimant une admiration amoureuse. Au-dessous est placée la photographie de la vedette, empruntée de toute évidence à l'un de ses sites.

(3) Dans sa deuxième réplique ci-dessous, l'admiratrice, qui communique sur Twitter (adresse : JoeSaderOfficial) exprime son « amour » pour l'acteur. De telles déclarations adressées à des vedettes sont très courantes, tant en France qu'au Proche-Orient, comme le montrent les phone-in analysés par V. Traverso (2006).

30 *Réponse de Joe Sader*

Traduction de l'anglais : « Merci ma douce (sweety) », réponse suivie de deux émoticônes « baiser » et d'une émoticône « fleur »

Commentaire : La réponse est conventionnelle et polie.

31 Dernière réponse de la 1^{re} interlocutrice, @mmm

Traduction de l'arabe dialectal libanais noté en arabizi : « Joe tu es magnifique, je t'aime » - « Joe, I'm mad for you, I love you ». Réponse suivie de l'émoticône « cœur » répétée trois fois.

Commentaire : L'expression *betjannen*, mot-à-mot « tu rends fou » (comparer avec, en arabe dialectal tunisien *bethabbel*, qui a le même sens) signifie ici un compliment appuyé, d'où la traduction proposée ci-dessus. Comme nous l'avons indiqué, ces déclarations d'amour à une vedette sont beaucoup plus conventionnelles qu'il n'y paraît à première vue.

32 De la même manière que pour la conjuration du mauvais œil (1^{er} mot de l'échange), l'arabe dialectal permet de dire les choses dans le cadre de la culture quotidienne des locuteurs. La transcription est en arabizi, et non en caractères arabes.

3.4. Exemple n°4, arabe libanais

Figure 7 : Reproduction anonymée du quatrième extrait

33 **Observations sur les données**

1^{er} interlocuteur, @nnn

Traduction : « Je ne vais pas mettre dans ma Biographie SSS [nom de l'Université], ça veut dire, je ne le mettrai pas » – « I'm not going to put in my Bio SSS [name of University], that is, I am not ». Suit une image de la notification reçue sur son téléphone, indiquant que son université « le suit ».

Commentaire : Tout l'échange est en arabe dialectal libanais et en arabizi, avec un seul emprunt à l'anglais, le mot Bio (noté avec une majuscule, s'agissant du CV « officiel » sur Internet), le mot parking (utilisé par @ooo) étant entré de longue date dans ce parler. Le ton de cette réplique indique que l'interlocuteur est en colère, en raison sans doute de la notification qu'il reproduit.

34 2nd interlocuteur, @ooo

Traduction : « Il y a maintenant une nouvelle application de l'université, qu'ils ont mise en route. Ils nous rendent fous avec ça, même sur le parking, ils nous envoient des pubs » – « There's now a new app from the university, in use. They madden us with it, even in the parking they send us adverts. »

Commentaire: Il est fait usage des grands caractères, ce qui indique une volonté d'expressivité. L'interlocuteur attire l'attention de son ami sur ce qui se produit dans leur université privée. L'expression est entièrement en dialectal noté en arabizi.

35 Réponse de @nnn

Il écrit : « hahaha ».

Commentaire : Les deux camarades d'université s'expriment en arabe dialectal libanais et en arabizi comme s'ils se parlaient. On est donc devant une interaction écrite, avec réaction quasi immédiate, analogue en partie à ce qu'on observe à l'oral.

4. Conclusion : vers une définition élargie de l'arabizi en interaction dans un contexte de communication multiforme

- 36 (1) La définition de l'arabizi comporte deux aspects : « qu'est-ce que l'arabizi ? » d'une part et « comment fonctionne l'arabizi ? » de l'autre. Le travail sur corpus conduit, comme on a pu le voir, à mettre l'accent sur le second.
- 37 (2) La littérature sur l'arabizi remonte, on l'a vu, à un peu moins d'une décennie. On lui doit un ensemble d'observations sur la transcription en caractères latins de l'arabizi.
- 38 (3) Mais les travaux cités, au-delà des aspects linguistiques, s'interrogent sur la libération de la parole sur les réseaux sociaux. Par exemple, ci-dessus, nous avons mis « interlocuteur » au masculin, en supposant qu'il s'agit d'hommes, au vu des pseudonymes ; or l'anonymat de ceux-ci permet toutes sortes de combinaisons s'inscrivant dans une libération de la parole entre personnes de sexe différents.
- 39 (4) Ce travail, inscrit dans le projet PARADE et dans ses analyses de corpus, montre que l'arabizi ne peut se réduire à des transcriptions, même si celles-ci constituent un aspect important de l'analyse. Dans nos corpus, nous avons observé la complexité des données « para-textuelles » : rôle de l'image ; rôle expressif des émoticônes et de la dimension des caractères.
- 40 (5) Le contexte sémantique et interactionnel des échanges en arabizi constitue un autre aspect essentiel : l'interprétation de nombreux détails des échanges provient de données contextuelles au sens large, y compris, comme on l'a vu, de données conversationnelles : il s'agit en fait d'échanges en interaction, ce qui suppose la mise en œuvre d'analyses fines.
- 41 (6) Il s'ajoute à cela la présence de données linguistiques plurielles : polyglossie à l'intérieur de la langue arabe considérée dans la multiplicité de ses usages (y compris à l'écrit, avec l'arabizi d'une part et l'usage de caractères arabes de l'autre), et alternance codique (*code-switching*) avec l'anglais ou le français.
- 42 (7) La présence de ces langues mises à part, les données indiquent que la communication en arabizi relève de l'ensemble complexe des usages inscrits dans la polyglossie de l'arabe, ce que donnent à voir les phénomènes relevant des arabes moyens ou l'alternance, dans la transcription entre l'arabizi et des énoncés en dialecte transcrits en caractères arabes. L'arabizi n'est donc pas, comme certains ont pu le penser, une « langue en soi » : il n'y pas de signes de créolisation ou de « métissage » linguistique inscrit dans un système. Les données montrent clairement que l'on n'y fait rien d'autre que ce que l'on fait en parlant.
- 43 (8) Notre enquête préliminaire sur corpus permet à la fois d'affiner la définition de la communication en arabizi, et de montrer la difficulté du travail sur des données aussi complexes (que faire par exemple des émoticônes ?). Des décisions devront être prises quant au niveau de granularité auquel l'analyse de corpus ou les traitements automatisés devront s'attacher. Mais pour la dialectologie arabe, un nouveau domaine, avec des types de corpus inconnus jusqu'à la dernière décennie, s'ouvre à la recherche.

BIBLIOGRAPHIE

- Aboeazz, M. 2009.. Latinised Arabic and connections to bilingual ability. *Lancaster University Postgraduate Conference in Linguistics & Language Teaching*. Department of Linguistics and English Language Lancaster University .
- Al-Khatib, M. & Sabbah, E. 2008. Language Choice in Mobile Text Messages among Jordanian. *SKY Journal of Linguistics* 21.
- Bahrainwala, L. .2011. You say Hello, I say Mar7aba: Exploring the digi-speak that powered the Arab revolution. Michigan State University.
- Bianchi, R. .2012. 3arabizi – When Local Arabic Meets Global English. *Acta Linguistica Asiatica* 2.
- Denis, A., Fréard, D., Détienne, F., Baker, M., & Barcellini, F. 2012. Détection de conflits dans les communautés épistémiques en ligne. *Actes de la Conférence sur le Traitement Automatique des Langues Naturelles*. Grenoble.
- Dichy, J. 2010. La polyglossie de l'arabe, illustrée par deux corpus d'époques et de natures différentes : un échange radiophonique syrien et un conte des Mille et Une Nuit. M. Bozdemir et L.-J. Calvet (éds.). *Les politiques linguistiques en Méditerranée*., Paris : H. Champion.
- Elmahdy, M., Gruhn, R., & Minker, W. 2012. *Novel Techniques for Dialectal Speech Recognition*. Springer Ed.
- Gonzales-Quijano, Y. 2012a. *Arabités numériques. Le printemps du Web arabe*, Arles : Actes Sud.
- Gonzales-Quijano, Y. 2012b. « Arabizi, le printemps linguistique. » From *Le Temps*: http://www.letemps.ch/Facet/print/Uuid/9d166cb8-1227-11e2-ae93-607a56226934/Arabizi_le_printemps_linguistique.
- peerreach.com. 2013. *4 ways how Twitter can keep growing*. From <http://blog.peerreach.com/2013/11/4-ways-how-twitter-can-keep-growing/>
- Randa, M., Mona, F., Nariman Elshamly, E. & Nady, A.-G. 2011. Summary of Arabizi or Romanization: The dilemma of writing Arabic texts. *Jil Jadid Conference*. Austin,; University of Texas.
- Traverso, V. 2006. *Des échanges ordinaires à Damas: aspects de l'interaction en arabe*. Lyon: Presses Universitaires de Lyon/Damas : Institut Français du Proche Orient .
- Yaghan, M. A. 2008. "Arabizi": A Contemporary Style of Arabic Slang. *Design Issues*.

RÉSUMÉS

Cette communication, inscrite dans le projet de recherche PARADE (partenaires : SYSTRAN, Laboratoire ICAR et ARADIC Editions), porte sur des analyses de copus en *arabizi*, qui correspond à l'usage sur Internet, réseaux sociaux ou SMS de l'arabe noté en caractères latins selon des modes de transcriptions issus de l'usage. Après un bref rappel de l'expansion d'Internet dans le monde arabe, de la littérature sur l'arabizi et des transcriptions qu'on y rencontre, les auteur(e)s procèdent à l'analyse de quatre brefs échanges sur Twitter. Deux concernent des internautes marocain(e)s et deux autres, des internautes libanais(es). Protégées par des pseudonymes, ces personnes s'expriment librement, à la fois quant au contenu et quant aux formes linguistiques qu'elles choisissent. Les données sont complexes. On y trouve : (1) des textes en arabizi,

transcrivant le dialecte de l'auteur(e) ; (2) des textes en caractères arabes dans l'un des dialectes, l'un des arabes moyens ou en arabe littéraire moderne ; (3) de l'anglais ou du français (voire de l'espagnol ou de l'italien), soit en tant que tels (alternance codique), soit sous forme d'emprunts ; (4) des émoticônes qui sont parfois répétées (itération) ; (5) un usage de caractères de grande dimension, à côté des caractères habituels, ce qui correspond à une fonction expressive ; (6) des images. Les échanges relèvent la plupart du temps d'une analyse interactionnelle, ce qui rend leur interprétation fortement dépendante du contexte. Cette première intervention montre que la question fondamentale – pour un linguiste – est, au-delà de « qu'est-ce l'abizi ? », celle de savoir « comment fonctionne l'arabizi ? »

INDEX

Mots-clés : arabizi, transcription, interaction écrite, polyglossie, arabes moyens, alternance codique, corpus complexes, réseaux sociaux, émoticônes

AUTEURS

JOSEPH DICHY

Professeur de linguistique arabe à Lyon et Laboratoire ICAR (UMR 5191, CNRS/Lyon 2, ENS-Lyon)
<http://.icar.univ-lyon2.fr>
Joseph.Dichy@yahoo.fr

FATIMA-ZAHRA DRISS

Aradic éditions et Laboratoire ICAR (UMR 5191, CNRS/Lyon 2, ENS-Lyon) <http://.icar.univ-lyon2.fr>
fatima-zahra.driss@aradic.fr

HASSIBA LAGRAF

Université Lumière-Lyon 2 et Laboratoire ICAR (UMR 5191, CNRS/Lyon 2, ENS-Lyon)
<http://.icar.univ-lyon2.fr>
hassibalg86@yahoo.fr

DJAMEL MOSTEFA

SYSTRAN SA - R&D, www.systran.com
djamel.mostefa@systrangroup.com