

Constitution d'un Thésaurus de la poésie arabe¹ Autour du *dīwān* de Jamīl Buṭayna

Georges Bohas, Djamel Eddine Kouloughli,
Maria Mouhieddine²

Depuis les temps les plus anciens, on considère la poésie comme le *dīwān* des Arabes. La constitution de corpus poétiques pose en partie des problèmes analogues à celle d'un corpus de prose littéraire, et en partie des problèmes spécifiques.

Concernant les problèmes identiques, citons d'abord celui de la sélection d'œuvres représentatives, pour chaque grande période de l'histoire littéraire de la langue arabe : poésie ancienne, poésie omeyyade, abbasside (avec peut-être des sous-périodes), moderne ; et aussi d'œuvres représentatives pour chaque genre, que ce soit du point de vue métrique – *qaṣīda*, *rajaz*, etc. – ou du point de vue thématique – poésie bédouine, amoureuse, bachique, politique, etc.

Toujours dans le domaine des problèmes identiques, il y a la nécessité (plus impérieuse encore que pour la prose) d'une vocalisation intégrale des textes. Dans une communication au colloque *Métrique arabe : état des lieux et perspectives* (Damas, 27-28 avril 2007), D. E. Kouloughli a montré que cette vocalisation devait s'effectuer sur un texte en transcription :

Le fondement du problème est que dans l'usage usuel de l'écriture arabe, les utilisateurs ont recours à toutes sortes de « simplifications », comme l'utilisation de graphèmes complexes combinant deux unités (*lām-alif*) et l'économie de certains graphèmes (notamment la *šadda* et surtout les voyelles brèves). En principe, dans la saisie soignée d'un texte, comme ce devrait être le cas pour un texte poétique que l'on veut soumettre à une analyse métrique, les éditeurs « font attention » et dotent le texte de tous les indices graphiques nécessaires à une lecture intégrale. Mais, en fait, les habitudes d'économie graphique se révèlent plus fortes que les meilleures résolutions, en sorte que pratiquement aucun texte saisi en graphie arabe n'est totalement exempt de « raccourcis » graphiques qui vont s'avérer autant de causes d'erreur pour l'algorithme d'analyse syllabique.³

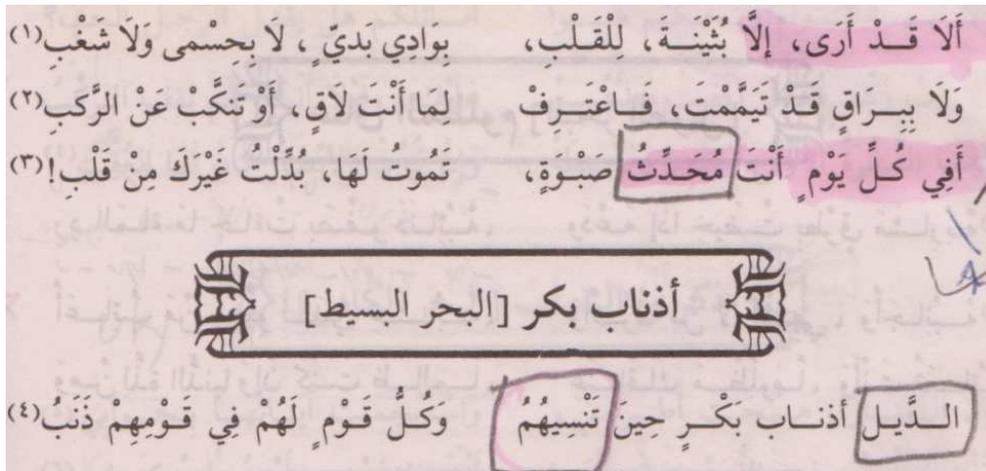
1. Cet article est le fruit d'une réflexion commune menée autour de la thèse de Maria Mouhieddine soutenue en novembre 2011, dirigée par G. Bohas et D. E. Kouloughli, *Aspects formels du dīwān de Jamīl Buṭayna*.

2. G. Bohas : ENS Lyon ; D. E. Kouloughli : CNRS ; M. Mouhieddine : ENS Lyon.

3. Kouloughli, 2010, p. 22-23.

Passons maintenant aux problèmes différents. Si l'on veut soumettre les textes à une analyse métrique, il faut déjà s'assurer de la correction de la saisie. En effet, rien ne nous dit que l'éditeur du texte soit compétent.

Si l'on considère la page suivante⁴ :



on constate qu'il faut corriger les mots encadrés :

أَفِي كُلِّ يَوْمٍ أَنْتَ مُحَدِّثٌ صَبْوَةً *** تَمُوتُ لَهَا بَدَّلْتُ غَيْرَكَ مِنْ قَلْبِ
الدَّيْلُ أَذْنَابُ بَكْرٍ حِينَ تَنْسِيهِمْ *** وَكُلُّ قَوْمٍ لَهُمْ فِي قَوْمِهِمْ ذَنْبٌ

Il est clair que *muḥaddiṭu* à la place de *muḥdiṭu* et *tansiyhumu* à la place de *tansibuhum* changent complètement l'analyse métrique. Quant à *al-daylu* à la place de *al-diylu* (nom d'une tribu), cela montre simplement l'inculture de l'éditeur. En fait, préalablement à tout traitement, il faudrait faire corriger le document par un très bon connaisseur de la langue arabe.

Dans ce domaine aussi notre équipe a déjà quelques réalisations à son actif. Pour la poésie ancienne ont été saisis, pour la plupart par ou sous la direction de D. E. Kouloughli, les *Mu'allaqāt*, le *dīwān d'al-Aḥṭal*, celui de *Jamīl Buṭayna*, celui de *al-Šarīf al-Raḍī* (l'édition papier de *Dār Ṣādir* compte quelque mille deux cents pages, ce qui donne une idée de l'ampleur de la tâche accomplie), et, concernant la poésie moderne, on dispose d'une partie du *dīwān* d'Adūnīs, d'un recueil de Unṣī Al-Ḥājj intitulé *Lan*, ainsi que de nombreux extraits de l'œuvre de Nizār Qabbānī⁵.

En ce qui concerne les outils de traitement, D. E. Kouloughli a réalisé (sans doute le seul) analyseur syllabique automatique de vers arabes. Nous allons décrire rapidement comment on le met en œuvre à partir de l'étude du *dīwān* de *Jamīl Buṭayna*⁶.

4. Extrait de l'édition de Maḥdī Muḥammad Nāṣir al-Dīn.

5. Nous nous limitons ici à la poésie classique.

6. Mouhieddine, 2011, est la première thèse réalisée dans ce cadre.

La première étape est de faire saisir le texte sans les voyelles. L'étape suivante consiste à convertir le texte arabe en transcription par le système TRS⁷ élaboré par D. E. Kouloughli. Cette conversion est effectuée automatiquement. Voici alors ce que l'on obtient pour le début de la première pièce de vers :

mn elxfret elbyD 'xLS lwnhe // tleHy & dwN lm yjd me y & ybhe
 fine mzn# byn elsmekyn 'wmDt // mn elnwr cm est & rDthe jnwbhe
 b'Hsn mnhe ywm qelt w & ndne // mn elnes 'wbe\$uN yxef \$gwbhe
 t & eyyt festgnyt & ne bgyrne // 'ilY ywm ylqY kl nfsiN Hbybhe
 wddt wle tgnY elwded# 'nhe // nSyby mn eldnye w'ny nSybhe

Il faut alors procéder à la vocalisation intégrale, ce qui donne :

min elxafaaraati elbiyDi 'uxliSa lawnuhaa // tulaaHiy & aduw*aN lam yajid
 maa ya & iybuhaa
 famaa muzna#uN bayna elsaamaakayni 'awmaDat // min elnuwri cum*a esta
 & raDathaa januwbuhaa
 bi'aHsana minhaa yawma qaalat wa & indanaa // min elnaasi 'awbaa\$uN
 yuxaafu \$uguwbuhaa
 ta & aayayta faestagnayta & an*aa bigayrinaa // 'ilaY yawmi yalqaY kul*a
 nafsIN Habiybuhaa
 wadidtu walaa tugniy elwidaada#u 'an*ahaa // naSiybiy min eldunya
 wa'an*iy naSiybuhaa

Les erreurs de vocalisation sont surlignées en bleu. Le problème est, à ce niveau, de contrôler les erreurs, surlignées ci-dessous en jaune.

1. PREMIER CONTRÔLE

L'analyse syllabique, réalisée par le logiciel Xaliyl, détecte toutes les erreurs de saisie qui ont pour effet de violer les règles de structure syllabique. En voici quelques exemples typiques.

1.1. MANQUE D'UNE VOYELLE

wamaa wajadat wajdiy bihaa 'um*u waaHidiN// walaa **wajad elnahdiy***u
 wajdiy & alaY hindi
 wamaa wajadat wajdiy bihaa 'um*u waaHidiN
 +wa + maa + wa + ja + dat + waj + diy + bi + haa + 'um + mu + waa + Hi + din
 CV CVV CV CV CVC CVC CVC CV CVV CVC CV CVV CV CVC
 v - v v - - - v - - v - v -
 walaa **wajad el**nahdiy*u wajdiy & alaY hindi
 ERREUR

laqad fuD*ilat HusnaN & alaY elnaasi **miclmaa**// & alaY 'alfi \$ahriN fuD*ilat
 layla#u elqadri
 laqad fuD*ilat HusnaN & alaY elnaasi mi**clmaa**
 ERREUR

7. En annexe, le système de transcription TRS.

&alaY 'alfi \$ahriN fuD*ilat layla#u elqadriy
 +&a + laa + 'al + fi + \$ah + rin + fuD + Di + lat + lay + la + tul + qad + riy
 CV CVV CVC CV CVC CVC CV CVC CVC CV CVC CVC CVC
 v - - v - - - v - - v - - -

1.2. INSERTION D'UN ESPACE INDÛ

wayan\$uru sir*aN fiy elSadiyqi wagayrahu// ya & iz *u & alaynaa na\$ruhu
 Hiyna yun\$aru
 wayan\$uru sir*aN fiy elSadiyqi wagayrahuw
 +wa + yan + \$u + ru + sir + ran + fil + Sa + diy + qi + wa + gay + ra + huw
 CV CVC CV CV CVC CVC CVC CV CVC CV CV CVC CV CVC
 v - v v - - - v - v v - v -
 ya & iz *u & alaynaa na\$ruhu Hiyna yun\$aruw
 ERREUR

wakun*aa jamiy & aN qabla 'an yaZhara elnawaY// bi' an & ami Haalay
 gibTa#iN wasuruwri
 wakun*aa jamiy & aN qabla 'an yaZhara elnawaY
 +wa + kun + naa + ja + miy + & an + qab + la + 'an + yaZ + ha + ral + na + waa
 CV CVC CVV CV CVC CVC CVC CV CVC CVC CV CVC CV CVV
 v - - v - - - v - - v - v -
 bi' an & ami Haalay gibTa#iN wasuruwriy
 ERREUR

1.3. MAJUSCULE POUR MINUSCULE

ma & aarifu lilxawdi el*atiy qultu 'ajmiliy// 'ilayhaa faqad 'aSfayti bilwud*I
 'ajma & a
 ma & aarifu lilxawdi el*atiy qultu 'ajmiliy
 +ma + & aa + ri + fu + lil + xaw + dil + la + tiy + qul + tu + 'aj + mi + liy
 CV CVV CV CV CVC CVC CVC CV CVC CVC CV CVC CV CVC
 v - v v - - - v - - v - v -
 'ilayhaa faqad 'aSfayti bilwud*I 'ajma & aa
 ERREUR

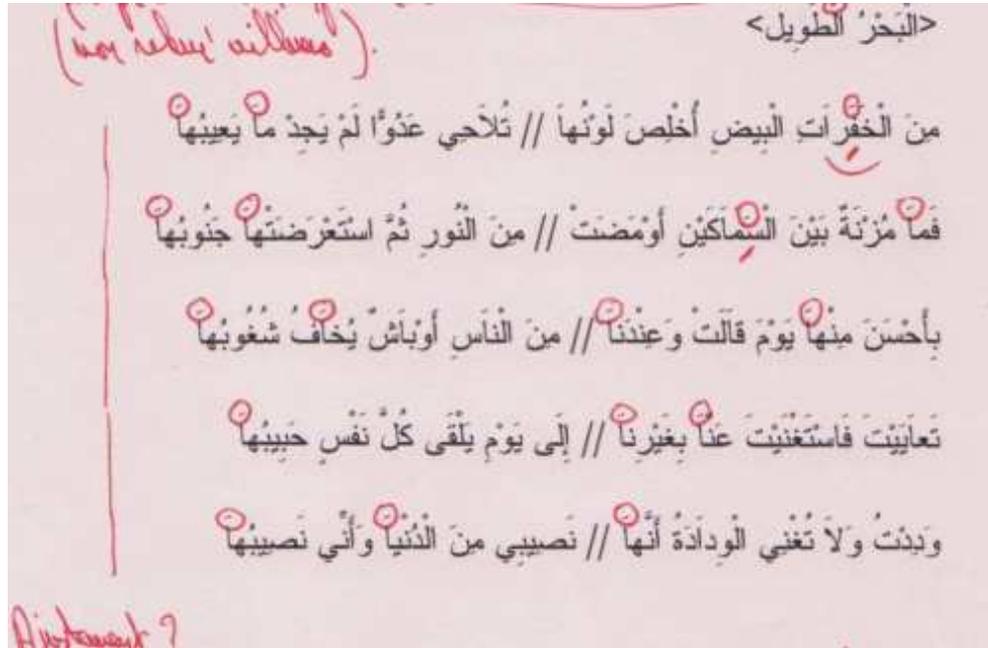
2. DEUXIÈME CONTRÔLE

Mais revenons à nos deux erreurs du début :

min elxafaaraati elbiyDi 'uxliSa lawnuhaa// tulaaHiy & aduw*aN lam yajid
 maa ya & iybuhaa
 famaa muzna#uN bayna elsaamaakayni 'awmaDat// min elnuwri cum*a esta
 & raDathaa januwbuhaa

Un *a* à la place d'un *i* n'a pas de conséquence sur l'analyse syllabique, donc l'erreur n'est pas détectée par l'analyseur syllabique. En d'autres termes, elle échappe au système d'analyse syllabique. Il faut un nouveau contrôle par un

spécialiste de l'arabe. À cet effet, on repasse le texte saisi en transcription en arabe et on le donne à un *mudaqqiq*⁸ :



Comme on peut le constater, *al-xafarāt* au lieu d'*al-xafirāt* n'a pas échappé au *mudaqqiq*, de même *al-samakayn* au lieu d'*al-simākayn*.

Il reste à enter les corrections pour avoir ce que nous appellerons un document source correct :

min elxafiiraati elbiyDi 'uxliSa lawnuhaa// tulaaHiy & aduw*aN lam yajid
 maa ya & iybuhaa
 famaa muzna#uN bayna elsimaakayni 'awmaDat// min elnuwri cum*a esta
 & raDathaa januwbuhaa
 bi'aHsana minhaa yawma qaalat wa & indanaa// min elnaasi 'awbaa\$uN
 yuxaafu \$uguwbuhaa
 ta & aayayta faestagnayta & an*aa bigayrinaa// 'ilaY yawmi yalqaY kul*a
 nafsiN Habiybuhaa
 wadidtu walaa tugnuy elwidaada#u 'an*ahaa// naSiybiy min eldunya
 wa'an*iy naSiybuhaa

Et pour revenir aux fautes de l'éditeur repérées au début :

'afiy kul*i yawmiN 'anta muHdicu Sabwa#iN// tamuwtu lahaa bud*iltu
 gayraka min qalbi
 eldiylu 'a@naabu bakriN Hiyna tansibuhum// wakul*u qawmiN lahum fiy
 qawmihim @anabu

On a donc maintenant un document source correct.

8. Ici, Ahyaf Sinno, de l'Université Saint-Joseph de Beyrouth.

Bifurcation

À partir d'un texte vocalisé correctement, on doit prendre une décision : Veut-on réaliser l'édition en arabe du texte ou un document que l'on soumettra à l'analyse métrique. En effet, dans l'analyse métrique doivent figurer des informations qui ne sont pas reportées dans le texte édité. Ainsi, pour l'analyse métrique, le pronom deuxième personne du pluriel peut être écrit *kumuw*, mais dans les conventions d'édition en arabe on l'écrit كُمْ et non كُمو. De même, le pronom première personne *'anaa* doit parfois être analysé *'ana* : il n'est pas possible cependant de l'écrire autrement que أنا. Il faut bifurquer :

- vers l'analyse métrique (3.) ou
- vers l'édition en arabe (4.).

3. ANALYSE MÉTRIQUE

Supposons donc que toutes les ERREURS de saisie aient été corrigées, et que nous disposions d'un texte analysé syllabiquement. C'est à ce moment qu'intervient l'analyse métrique proprement dite.

3.1. DES CAS SIMPLES

Considérons le vers suivant, sur le mètre *mutaqārib* dont le profil métrique est⁹ :

[v -] x	[v -] x	[v -] x	([v -] -)	§	[v -] x	[v -] x	[v -] x	([v -] -)
wa'i@ 'anaa	'agyadu	gaD*u	el\$abaabi//		'ajur*u	elridaa'a	ma &	a elmiIzari

Considérons le premier hémistiche :

wa'i@ 'anaa 'agyadu gaD*u el\$abaabiy
CV CVC CV CVV CVC CV CV CVC CVC CV CVV CVC

v - v - v - v - - - : on voit que la longue surlignée contredit le modèle :

[v -] x	[v -] x	[v -] x	([v -] -)	§	[v -] x	[v -] x	[v -] x	([v -] -)
---------	---------	---------	-----------	---	---------	---------	---------	-----------

La position surlignée en bleu doit être brève. Ici il n'y a pas d'erreur de saisie, c'est la métrique qui nous dit qu'il faut lire *'ana* et non *'anaa* :

wa'i@ 'ana 'agyadu gaD*u el\$abaabiy

pour obtenir :

v - v//v - v//v - -//v - - qui est conforme au modèle.

Autre exemple concernant le *mutaqārib* :

fagay*ara @alika maa ta & lamiyna// tagay*ura @aa elzamani elmunkari

Considérons le premier hémistiche :

9. Le x se réalise en une syllabe brève ou longue ; voir Bohas, 2010.

fagay*ara @alika maa ta & lamiynaa
 CV CVC CV CV CV CV CV CVV CVC CV CVC CVV
 v - v v v v v - - v - -

On observe une suite de 5 brèves, ce qui est impossible en métrique arabe. Il faut optimiser en allongeant le a de @alika :

fagay*ara @aalika maa ta & lamiynaa

ce qui donne bien un *mutaqārib* correct :

v - v // v - v // v - - // v - -

Abordons un autre cas dans un *ṭawīl* dont le profil métrique est :

[v -]	x	[v -]	x x	[v -]	x	[v -]	x -
-------	---	-------	-----	-------	---	-------	-----

fa'in lam tahjiniy @aata yawmiN fa'in*ahuw
 CV CVC CVC CVC CV CVC CVV CV CVC CVC CV CVC CV CVC
 v - - - v - - v - - v - v -
 sayakfiyka warqaa'u elsaraa#i SaduwHuw
 +sa + yak + fiy + ka + war + qaa + 'ul + sa + raa + ti + Sa + duw + Huw
 CV CVC CVC CV CVC CVV CVC CV CVV CV CV CVC CVC
 v - - v - - - v - v v - -

Dans le premier hémistiche, la séquence surlignée en bleu n'entre pas dans le profil du *ṭawīl*.

Ici, l'analyse métrique permet de détecter une erreur de saisie qui avait échappé au programme d'analyse syllabique ; il faut lire :

fa'in lam tahjiniy @aata yawmiN fa'in*ahuw
 +fa + 'in + lam + ta + hij + niy + @aa + ta + yaw + min + fa + 'in + na + huw
 CV CVC CVC CV CVC CVC CVV CV CVC CVC CV CVC CV CVC
 v - - // v - - - // v - - // v - v -
 sayakfiyka warqaa'u elsaraa#i SaduwHuw
 +sa + yak + fiy + ka + war + qaa + 'ul + sa + raa + ti + Sa + duw + Huw
 CV CVC CVC CV CVC CVV CVC CV CVV CV CV CVC CVC
 v - - // v - - - // v - v // v - -

ce qui donne bien un *ṭawīl* correct.

3.2. LES HAMZA

Restons dans le *ṭawīl* et examinons le vers suivant :

'i@aa fak*arat qaalat qad'adraktu wud*ahuw
 CV CVV CVC CV CVC CVV CVC CVC CVC CVC CV CVC CV CVC
 v - - v - - - - - v - v -
 wamaa Dar*aniy buxliy fakayfa 'ajuwduw
 +wa + maa + Dar + ra + niy + bux + liy + fa + kay + fa + 'a + juw + duw
 CV CVV CVC CV CVC CVC CVC CV CVC CV CV CVC CVC
 v - - v - - - v - v v - -

Il est bien évident que la séquence de longues du premier hémistiche ne réalise pas le profil du *ṭawīl*. La solution est de transformer la *hamza al-qaṭ'* en *waṣl* :

'i@aa fak*arat qaalat qad adraktu wud*ahuw
 CV CVV CVC CV CVC CVV CVC CV CVC CV CVC CV CVC
 v - -//v - - -//v - -//v - v -

La saisie était correcte, mais c'est la métrique qui nous amène à modifier le texte en supprimant la *hamza al-qaṭ'* pour obtenir un *ṭawīl* conforme au profil.

3.3. LES PRONOMS

Passons au *basīṭ* dont le profil est :

x x [v -]	x [v -]	x x [v -]	x [v -]	x x [v -]	x [v -]	x x [v -]	x [v -]
-----------	---------	-----------	---------	-----------	---------	-----------	---------

qad kuntu & ankum ba & iyda eldaari mugtaribaN
 CVC CVC CV CVC CVC CV CVC CVC CVV CV CVC CV CV CVC
 - - v - - v - - - v - v v -
 Hat*aY da & aaniy liHayniy **minkum daa & iy**
 CVC CVV CV CVV CVC CV CVC CVC CVC CVC CVV CVC
 - - v - - v - - - - - - - -

Une pareille séquence de longues est impossible. Il faut donc lire **minkumuW** :

Hat*aY da & aaniy liHayniy **minkumuW** daa & iy
 +Hat + taa + da + & aa + niy + li + Hay + niy + min + ku + muw + daa + & iy
 CVC CVV CV CVV CVC CV CVC CVC CVC CV CVC CVV CVC
 - - v - - v - - - v - - -
 ce qui donne bien un *basīṭ* correct :
 - - v -// - v -// - - v -// - -

Il en ira de même avec *hum/humuW*.

Une fois qu'on a optimalisé le texte de la sorte, l'analyse syllabique est métriquement correcte. Voici le dernier état :

qad kuntu & ankum ba & iyda eldaari mugtaribaN// Hat*aY da & aaniy
 liHayniy minkumuW daa & i
 qad kuntu & ankum ba & iyda eldaari mugtaribaN
 +qad + kun + tu + & an + kum + ba + & iy + dal + daa + ri + mug + ta + ri + ban
 CVC CVC CV CVC CVC CV CVC CVC CVV CV CVC CV CV CVC
 - v - - v - - - v - v v -
 Hat*aY da & aaniy liHayniy **minkumuW** daa & iy
 +Hat + taa + da + & aa + niy + li + Hay + niy + min + ku + muw + daa + & iy
 CVC CVV CV CVV CVC CV CVC CVC CVC CV CVC CVV CVC
 - - v - - v - - - **v** - - -

et on peut constater que l'analyse métrique est parfaite :

- - v -// - v -// - - v -//v v -
 - - v -// - v -// - - **v** -// - -

Quand on travaille sur l'analyse syllabique, il faut donc revenir au fichier source, le corriger et réanalyser pour obtenir l'analyse correcte.

3.4. UN CAS PLUS SOPHISTIQUE

Reprenons le profil du *ṭawīl* :

[v -]	x	[v -]	x x	[v -]	x	[v -]	x -
-------	---	-------	-----	-------	---	-------	-----

Depuis longtemps, on s'est aperçu que les x x du deuxième pied de chaque hémistiche se réalisaient presque toujours en longues : [v -] ——. Elles constituent ce que nous avons appelé une zone de stabilité que nous allons surligner en gris. Pour constituer cette zone, les éléments variables concernés seront réalisés en longues. Nous noterons cela par PREFLONG (préférer les longues). Dans le système khalīlien on formulerait la chose en disant que l'on évite au maximum les *zihāfāt*.

Considérons ce vers :

wa'in*iy lamu\$taaquN 'ilaY riyHi jaybihaa// kamaa e\$taaqa 'idriysu 'ilaY
jan*a#i elxuldi

Cela donnerait :

v - - v - - - v - - v - v - v - - v - - v v - - v - - -

C'est la réalisation de la zone de stabilité qui amène la nûnation de 'idriysu, qui, en tant que nom propre étranger, est *mamnū' min al-ṣarf*, mais comme la nûnation des *mamnū' min al-ṣarf* est possible en poésie, PREFLONG impose la lecture 'idriysuN.

wa'in*iy lamu\$taaquN 'ilaY riyHi jaybihaa// kamaa e\$taaqa 'idriysuN 'ilaY
jan*a#i elxuldi

Ce qui donne :

v - - // v - - - // v - - // v - v - v - - // v - - // v - - // v - - -

avec respect de la zone de stabilité¹⁰.

4. ÉDITION DU TEXTE EN ARABE

Revenons à notre document de base, dûment corrigé :

min elxafiiraati elbiyDi 'uxliSa lawnuhaa// tulaaHiy & aduw*aN lam yajid
maa ya & iybuhaa
famaa muzna#uN bayna elsiimaakayni 'awmaDat// min elnuwri cum*a esta
& raDathaa januwbuhaa
bi'aHsana minhaa yawma qaalat wa & indanaa// min elnaasi 'awbaa\$uN
yuxaafu \$uguwbuhaa
ta & aayayta faestagnayta & an*aa bigayrinaa// 'ilaY yawmi yalqaY kul*a
nafsiN Habiybuhaa
wadidtu walaa tugniy elwidaada#u 'an*ahaa// naSiybiy min eldunyaa
wa'an*iy naSiybuhaa

10. Dans le *dīwān* de Jamīl, il n'y a qu'un cas où la zone de stabilité n'est pas respectée (voir Mouhieddine, 2011, p. 65).

Pour réaliser l'édition du texte en arabe, la première étape consiste à le convertir en écriture arabe¹¹ en ajoutant certaines modifications :

- 1) Mettre les *šadda* des solaires : **الْبَحْرُ الطَّوِيلُ**
 - 2) Normaliser l'écriture des *fatha alif*; à savoir, placer la *fatha* sur la consonne précédant le *alif*, comme dans : **فَمَا لَوْئَهَا** et **فَمَا**.
 - 3) Insérer les voyelles de liaison : **مِنَ النَّاسِ**.
- Ce qui donnera pour le début du premier poème du *dīwān* :

مِنَ الْخَفَرَاتِ الْبَيْضِ أَخْلَصَ لَوْئَهَا *** ثَلَاثِي عَدُوًّا لَمْ يَجِدْ مَا يَعِيبُهَا
فَمَا مَزْنَةٌ بَيْنَ السَّمَائِينَ أَوْمَضَتْ *** مِنَ الثُّورِ ثُمَّ اسْتَعْرَضَتْهَا جَنُوبُهَا
بِأَحْسَنَ مِنْهَا يَوْمَ قَالَتْ وَعَدَدْنَا *** مِنَ النَّاسِ أَوْبَاشٌ يُخَافُ شُعُوبُهَا
تَعَابَيْتَ فَاسْتَعْنَيْتَ عَنَّا بَعِيرَنَا *** إِلَى يَوْمٍ يَلْقَى كُلُّ نَفْسٍ حَبِيبُهَا
وَدِدْتُ وَلَا تُغْنِي الْوَدَادَةَ أَنَّهُ *** نَصِيبِي مِنَ الدُّنْيَا وَأَنِّي نَصِيبُهَا

Il reste un problème lié aux conventions graphiques arabes. En effet, si l'on contracte le *alif* et le *lām*, la voyelle intermédiaire présente dans la vocalisation systématique saute. Si l'on veut réaliser une édition intégralement vocalisée, on est donc, pour l'instant, obligé de maintenir la disjonction, ce qui paraît « peu esthétique » pour le lecteur arabophone non prévenu, puisqu'on obtient :

فَمَا ذَكَرَ الْخُلَانَ إِلَّا ذَكَرْتُهَا *** وَلَا الْبُخْلُ إِلَّا قُلْتُ سَوْفَ تَجُودُ
يُرْزَعُ فِيهَا الرِّيحُ كُلَّ عَشِيَّةٍ *** هَزِيمٌ بِسَلْفِ الرِّيحِ رَجِيعُ
وَأَنِّي لَأَسْتَبْكِي إِذَا ذَكَرَ الْهُوَى *** إِلَيْكَ وَأَنِّي مِنْ هَوَاكَ لَأَوْجِلُ
حَاوَلْتُ لِيَأْتِ حَبْلٌ وَصَالِكُمْ *** مَنِّي وَلَسْتُ وَإِنْ جَهْدَنْ بِفَاعِلِ

Cette particularité graphique se retrouve (par exemple sur le Web) pour la plupart des textes arabes systématiquement vocalisés. Cependant, si elle s'avère inacceptable pour le lecteur arabophone on pourra, bien sûr, convertir les séquences problématiques pour éditer un texte plus conforme à ses habitudes de lecture. Ainsi, le premier vers ci-dessus deviendrait :

فَمَا ذَكَرَ الْخُلَانَ إِلَّا ذَكَرْتُهَا *** وَلَا الْبُخْلُ إِلَّا قُلْتُ سَوْفَ تَجُودُ

5. CONCLUSION

À partir d'un *dīwān* systématiquement édité avec vocalisation intégrale, et analysé métriquement, nous espérons avoir montré comment on pourra constituer peu à peu un Thésaurus de la poésie arabe. L'exploitation d'un tel Thésaurus permettra non seulement de réaliser des éditions correctes, mais fera progresser aussi bien l'étude de la métrique arabe, celle de la poésie formulaire et, en fin de compte, participera au développement de l'histoire de la littérature arabe.

11. Cette conversion, comme la conversion inverse, s'effectue de façon automatique.

Références

- BOHAS G., 2010, « De la mesure en arabe : une description unifiée », *Bulletin d'études arabes*, n° 59, p. 33-60.
- KOULOUGHLI D. E., 2010, « Traitement automatique de la métrique arabe, réalisations et perspectives », *Bulletin d'études arabes*, n° 59, p. 17-31.
- MOUHIEDDINE M., 2011, *Aspects formels du dīwān de Jamīl Buṭayna*, Thèse de doctorat, ENS Lyon.
- Dīwān Jamīl Buṭayna*, édité par Maḥdī Muḥammad Nāṣir al-Dīn, 1993, Damas, Dār al-kutub al-‘ilmiyya.

Annexe

Table de conversion du système TRS

Transcription	Caractère arabe	Transcription	Caractère arabe
'	ء	&	ع
I	ئ	g	غ
U	ؤ	f	ف
e	ا	q	ق
Y	ى	k	ك
b	ب	l	ل
t	ت	m	م
#	ة	n	ن
c	ث	h	ه
j	ج	w	و
H	ح	y	ي
x	خ	*	ء
d	د	a	َ
@	ذ	i	ِ
r	ر	u	ُ
z	ز	aa	أ
s	س	iy	ي
\$	ش	uw	و
S	ص	aN	أ
D	ض	iN	ئ
T	ط	uN	ء
Z	ظ	aYN	ي