

Liste des figures

Figure 1 : Les moments de l'analyse lexicométrique.....	54
Figure 2 : La première page de l'édition de référence scannée d' <i>al-Ḥimtâ' wa-l-Muḥâna</i>	56
Figure 3 : Résultat de l'OCR avant apprentissage.....	56
Figure 4 : Même après apprentissage, le résultat de l'OCR est loin d'être satisfaisant.....	57
Figure 5 : Les étapes du dépouillement lexical.....	65
Figure 6 : Capture d'écran : écran d'accueil du segmenteur montrant le chemin du fichier à segmenter, le bouton du lancement de la segmentation et le champ affichant le n° du mot graphique eu cours de segmentation.....	72
Figure 7 : Capture d'écran : choix de segmentation à faire parmi deux propositions présentées par le segmenteur.....	73
Figure 8 : Apprentissage du segmenteur : segmentation manuelle des mots graphiques non analysés par le segmenteur et que ce dernier devra enregistrer...	74
Figure 9 : Fichier à segmenter avant de le soumettre au segmenteur.....	75
Figure 10 : Le fichier de sortie après segmentation.....	75
Figure 11 : Captures d'écran : Choix du lemme.....	87
Figure 12 : Hiérarchie de la classification des catégories lexicales.....	89
Figure 13 : Capture d'écran : insertion du code des verbes.....	92
Figure 14 : Capture d'écran : insertion du code des noms primitifs.....	93
Figure 15 : Capture d'écran : insertion du code des noms dérivés.....	93
Figure 16 : Capture d'écran : insertion du code des adjectifs.....	93
Figure 17 : Capture d'écran : insertion du code des mots-outils.....	94
Figure 18 : Capture d'écran : insertion du code des noms propres.....	94
Figure 19 : Capture d'écran : à la fin de l'opération de catégorisation chaque lemme est suivi du code de la catégorie lexicale correspondante.....	95
Figure 20 : Capture d'écran : fichier final de la catégorisation et où les codes ont remplacé les lemmes.....	96
Figure 21 : Répartition des mots graphiques à une seule ou à plusieurs unités lexicales.....	111
Figure 22 : Répartition de l'ensemble des mots graphiques selon le nombre d'unités les composant.....	111
Figure 23 : Répartition des mots graphiques multipartites.....	112
Figure 24 : Provenance des unités lexicales selon le type de mots graphiques...	112
Figure 25 : Coefficient de productivité des mots graphiques.....	114
Figure 26 : Ecran principal de Lexico3 après la création d'une nouvelle base.....	121
Figure 27 : Capture d'écran du fichier de sortie « <i>Lemm-.par</i> ».....	121
Figure 28 : Capture d'écran du fichier de sortie « <i>atrace.txt</i> ».....	122
Figure 29 : Extraits de 3 fichiers de sortie ".dic".....	122
Figure 30 : La barre d'outils de <i>Lexico3</i> dans laquelle figure la majorité des modules.....	123
Figure 31 : Concordance du lemme <i>مسألة</i> récupérée à partir de <i>Lexico3</i>	125

Figure 32 : Dictionnaire des segments répétées et TGen dans <i>Lexico3</i>	126
Figure 33 : Capture d'écran représentant les PCLC de notre corpus selon la partition Nuit.....	128
Figure 34 : Captures d'écran représentant les spécificités de la partie 08 pour la partition Nuit.....	129
Figure 35 : Capture d'écran représentant la carte des sections pour le lemme <i>قَالَ</i>	131
Figure 36 : Capture d'écran du dictionnaire de fréquences (ici hiérarchique).....	135
Figure 37 : Captures d'écran représentant les menus déroulants de <i>XLSTAT</i> et ses barres d'outils.....	138
Figure 38 : Représentation graphique en 3D faite avec <i>XLSTAT-3Dplot</i>	140
Figure 39 : Systèmes segmentés et systèmes non segmentés.....	200
Figure 40 : Les types de segmentation.....	201
Figure 41 : Schéma général du mot graphique en arabe.....	208
Figure 42 : Catégorisation classique des Verbes contenant les Verbes figés.....	263
Figure 43 : Catégorisation classique des Noms contenant les Noms-outils.....	264
Figure 44 : Organisation générale du lexique arabe dans une perspective de lexicométrie.....	274
Figure 45 : Capture d'écran de l'interface de saisie et de mise à jour des mots-outils de <i>DIINAR</i> : Choix de la catégorie.....	291
Figure 46 : Choix de la sous-catégorie.....	291
Figure 47 : Choix de la sous-catégorie de niveau 2.....	292
Figure 48 : Choix du mot-outil.....	292
Figure 49 : Le mot-outil avec les spécificateurs du niveau du mot qui lui sont associés.....	292
Figure 50 : Le mot-outil simple avec ses différentes agglutinations possibles....	293
Figure 51 : Catégorisation des mots-outils ayant servi à la constitution de la base de données des mots-outils de <i>DIINAR</i>	294
Figure 52 : Nombre de racines dans <i>DIINAR</i> selon R1.....	328
Figure 53 : Nombre de racines dans <i>Að-Niîâî</i> selon R1.....	335
Figure 54 : Nombre de racines dans <i>lisân al-ÝArab</i> selon R1.....	338
Figure 55 : Comparatif de la répartition des racines dans les trois dictionnaires selon R1.....	341
Figure 56 : Nombre de racines dans <i>Al-ðImtâÝ wa-l-Muðânasa</i> selon R1.....	347
Figure 57 : Diagramma de Pareto des racines dans <i>Al-ðImtâÝ wa-l-Muðânasa</i>	351
Figure 58 : Racines, formes et occurrences dans <i>Al-ðImtâÝ wa-l-Muðânasa</i>	353
Figure 59 : Diagramme de Pareto de la production des formes par groupe de racines selon R1 dans <i>Al-ðImtâÝ wa-l-Muðânasa</i>	356
Figure 60 : Diagramme de Pareto de la production des occurrences par groupe de racines selon R1 dans <i>Al-ðImtâÝ wa-l-Muðânasa</i>	360
Figure 61 : Productivité générale, en formes, des racines (groupées par R1): écart par rapport à la moyenne.....	370
Figure 62 : Productivité générale, en occurrences, des racines (groupées par R1): écart par rapport à la moyenne.....	370

Figure 63 : Productivité, en formes, positive, moyenne et négative des racines dans <i>Al-ḥImtâY wa-l-Muḥânasa</i>	375
Figure 64 : Productivité, en occurrences, faible et forte des racines dans <i>Al-ḥImtâY wa-l-Muḥânasa</i>	376
Figure 65 : Diagramme de Pareto de la production des formes par les racines dans <i>Al-ḥImtâY wa-l-Muḥânasa</i>	380
Figure 66 : Diagramme de Pareto de la fréquence des formes produites par les racines dans <i>Al-ḥImtâY wa-l-Muḥânasa</i>	382
Figure 67 : Nombre de racines dans <i>Les séances de HamÆânî</i> selon R1.....	385
Figure 68 : Nombre de racines dans <i>Al-ḥAdab al-Kabîr</i> selon R1.....	389
Figure 69 : Nombre de racines dans <i>Le Coran</i> selon R1.....	393
Figure 70 : Pourcentage des racines des trois corpus par rapport à DIINAR.....	399
Figure 71 : Pourcentage des racines des trois corpus par rapport à <i>lisân al-ÝArab</i>	399
Figure 72 : Pourcentage des racines des trois corpus par rapport à <i>Að-Ñîlâî</i>	400
Figure 73 : Pourcentage des racines des trois corpus par rapport à la moyenne théorique des dictionnaires.....	401
Figure 74 : Comparatif de la répartition des racines dans les trois corpus selon la 1 ^{ère} radicale.....	403
Figure 75 : Pourcentage des racines d' <i>Al-ḥAdab al-Kabîr</i> par rapport à DIINAR selon la 1 ^{ère} radicale.....	404
Figure 76 : Pourcentage des racines des <i>séances de HamÆânî</i> par rapport à DIINAR selon la 1 ^{ère} radicale.....	405
Figure 77 : Pourcentage des racines d' <i>Al-ḥImtâY wa-l-Muḥânasa</i> par rapport à DIINAR selon la 1 ^{ère} radicale.....	406
Figure 78 : Pourcentage des racines d' <i>Al-ḥAdab al-Kabîr</i> par rapport à <i>lisân al-ÝArab</i> selon la 1 ^{ère} radicale.....	409
Figure 79 : Pourcentage des racines des <i>séances de HamÆânî</i> par rapport à <i>lisân al-ÝArab</i> selon la 1 ^{ère} radicale.....	410
Figure 80 : Pourcentage des racines d' <i>Al-ḥImtâY wa-l-Muḥânasa</i> par rapport à <i>lisân al-ÝArab</i> selon la 1 ^{ère} radicale.....	411
Figure 81 : Pourcentage des racines d' <i>Al-ḥAdab al-Kabîr</i> par rapport à <i>Að-Ñîlâî</i> selon la 1 ^{ère} radicale.....	414
Figure 82 : Pourcentage des racines des <i>séances de HamÆânî</i> par rapport à <i>Að-Ñîlâî</i> selon la 1 ^{ère} radicale.....	415
Figure 83 : Pourcentage des racines d' <i>Al-ḥImtâY wa-l-Muḥânasa</i> par rapport à <i>Að-Ñîlâî</i> selon la 1 ^{ère} radicale.....	416
Figure 84 : Pourcentage des racines d' <i>Al-ḥAdab al-Kabîr</i> par rapport à la moyenne des dictionnaires selon la 1 ^{ère} radicale.....	421
Figure 85 : Pourcentage des racines des <i>séances de HamÆânî</i> par rapport à la moyenne des dictionnaires selon la 1 ^{ère} radicale.....	422

Figure 86 : Pourcentage des racines d' <i>Al-bImtâY wa-l-Muḥâḥana</i> par rapport à la moyenne des dictionnaires selon la 1 ^{ère} radicale.....	423
Figure 87 : L'écart, en pourcentage, des racines de chaque ouvrage par rapport à la base de données DIINAR.....	428
Figure 88 : L'écart, en pourcentage, du nombre des racines de chaque ouvrage par rapport à <i>lisân al-ÝArab</i>	429
Figure 89 : L'écart, en pourcentage, du nombre des racines de chaque ouvrage par rapport à <i>Aḏ-Ñiḥâf</i>	430
Figure 90 : Ecart du pourcentage des racines trilitères de chacun des trois corpus par rapport à la moyenne théorique des dictionnaires.....	432
Figure 91 : Ecart du pourcentage des racines quadrilitères de chacun des trois corpus par rapport à la moyenne théorique des dictionnaires.....	433
Figure 92 : Proportion entre racines trilitères et racines quadrilitères dans trois corpus poétiques modernes.....	436
Figure 93 : Nombre des racines dans <i>Ḥîḥâḥ al-bAfýâl al-ÝArabiyya</i> selon R1.....	439
Figure 94 : Diagramme comparatif des racines verbales entre <i>Le Coran</i> et <i>Ḥîḥâḥ al-bAfýâl al-ÝArabiyya</i>	441
Figure 95 : Pourcentage des racines verbales dans <i>Le Coran</i> par rapport aux racines de <i>Ḥîḥâḥ al-bAfýâl al-ÝArabiyya</i> selon la 1 ^{ère} radicale.....	442
Figure 96 : Etendue des <i>Nuits</i>	450
Figure 97 : Etendue des <i>Nuits</i> : distribution autour de l'étendue moyenne.....	450
Figure 98 : Etendue des <i>Nuits</i> et de leur vocabulaire.....	451
Figure 99 : Graphique de dispersion des rangs selon les valeurs de N et de V.....	454
Figure 100 : Relation entre N et V.....	455
Figure 101 : Répartition intertextuelle.....	459
Figure 102 : Répartition intertextuelle en pourcentage d'occurrences.....	460
Figure 103 : Répartition des versets selon les <i>Nuits</i>	461
Figure 104 : Répartition des vers selon les <i>Nuits</i>	462
Figure 105 : Ecart des vers par rapport à la moyenne.....	463
Figure 106 : Répartition générique en pourcentage d'occurrences.....	466
Figure 107 : informations contenues dans le TLE et servant au calcul des spécificités.....	468
Figure 108 : Classement de quelques nuits en fonction de la richesse lexicale selon la méthode des indices.....	487
Figure 109 : La richesse lexicale des <i>Nuits</i> d' <i>Al-bImtâY wa-l-Muḥâḥana</i> selon la formule de Guiraud.....	491
Figure 110 : Graphique de dispersion des rangs entre le classement selon N et celui selon la richesse lexicale.....	491
Figure 111 : La richesse lexicale des <i>Nuits</i> d' <i>Al-bImtâY wa-l-Muḥâḥana</i> selon la formule de Guiraud: Ecart par rapport à la richesse moyenne.....	492

Figure 112 : La richesse lexicale des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> selon l'indice W de Brunet.....	496
Figure 113 : La richesse lexicale des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> selon l'indice W de Brunet (valeurs ramenées à R).....	496
Figure 114 : Graphique de dispersion des rangs entre l'indice W et l'étendue N.....	497
Figure 115 : Richesse lexicale des <i>Nuits</i> selon l'indice W: Ecart par rapport à la richesse moyenne.....	498
Figure 116 : Le vocabulaire, réel et théorique, des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i>	504
Figure 117 : La richesse lexicale des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> selon la loi binomiale (méthode Muller).....	504
Figure 118 : Graphique de dispersion des rangs entre l'étendue des <i>Nuits</i> et la richesse lexicale selon la loi binomiale.....	506
Figure 119 : Richesse lexicale des <i>Nuits</i> selon la loi binomiale: Ecart par rapport à la richesse moyenne.....	507
Figure 120 : La richesse lexicale des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> selon l'indice V_m de Yule-Herdan.....	511
Figure 121 : Confirmation de l'influence de V1 sur l'indice V_m en comparant les deux diagrammes de la richesse lexicale, selon l'indice V_m et le quotient $V1/V$	512
Figure 122 : Graphique de dispersion des rangs entre l'étendue des <i>Nuits</i> et la richesse lexicale selon l'indice V_m	513
Figure 123 : Richesse lexicale des <i>Nuits</i> selon l'indice V_m : Ecart par rapport à la richesse moyenne.....	514
Figure 124 : Courbes des classements en fonction de la richesse lexicale des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> selon les différentes méthodes utilisées.....	518
Figure 125 : Représentation graphique, en 3 dimensions, du classement des nuits en fonction de la richesse lexicale selon les méthodes.....	521
Figure 126 : Analyse factorielle des variables latentes du classement des 15 nuits par les quatre méthodes de mesure de richesse lexicale : Représentations graphique des méthodes.....	523
Figure 127 : Analyse factorielle des variables latentes du classement des 15 nuits par les quatre méthodes de mesure de richesse lexicale : Représentations graphique des nuits.....	524
Figure 128 : La dispersion des nuits interprétée selon l'axe F1.....	526
Figure 129 : La dispersion des nuits interprétée selon l'axe F2.....	527
Figure 130 : Interprétation globale, selon les deux axes F1 et F2, de la représentation graphique de l'analyse factorielle des variables latentes.....	529
Figure 131 : Classement, en fonction de la richesse lexicale, des <i>Nuits</i> d' <i>Al-ḤimtâY wa-l-Muḥâna</i> résultant de la représentation graphique de l'analyse factorielle des variables latentes.....	531
Figure 132 : Accroissement réel du vocabulaire dans <i>Al-ḤimtâY wa-l-Muḥâna</i>	547

Figure 133 : Accroissement du vocabulaire d' <i>Al-Ḥimtâʿ wa-l-Muḥâna</i> par classe de fréquence.....	550
Figure 134 : Corrélation des rangs entre les classements selon l'accroissement lexical, l'étendue relative et l'ordre chronologique.....	557
Figure 135 : Accroissement cumulé du vocabulaire dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	558
Figure 136 : Accroissement réel et accroissement théorique du vocabulaire dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	559
Figure 137 : Lexicalité et fonctionnalité au niveau du corpus.....	566
Figure 138 : Lexicalité et fonctionnalité selon les <i>Nuits</i>	567
Figure 139 : Lexicalité et fonctionnalité des <i>Nuits</i> d' <i>Al-Ḥimtâʿ wa-l-Muḥâna</i> : Ecart par rapport à la moyenne.....	568
Figure 140 : Répartition des catégories lexicales au niveau du corpus entier....	570
Figure 141 : Répartition des Verbes dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	573
Figure 142 : Répartition des Verbes trilitères augmentés dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	573
Figure 143 : Répartition des Noms primitifs dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	575
Figure 144 : Répartition des Adjectifs dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	576
Figure 145 : Répartition des Noms dérivés dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	578
Figure 146 : Répartition des Particules dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	580
Figure 147 : Répartition des Noms-outils dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	583
Figure 148 : Répartition des Verbes fonctionnalisés dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	584
Figure 149 : Répartition des catégories lexicales selon les <i>Nuits</i> dans <i>Al-Ḥimtâʿ wa-l-Muḥâna</i>	586
Figure 150 : Répartition des Verbes selon les <i>Nuits</i>	588
Figure 151 : Répartition des Noms primitifs selon les <i>Nuits</i>	589
Figure 152 : Répartition des Adjectifs selon les <i>Nuits</i>	590
Figure 153 : Répartition des Noms dérivés selon les <i>Nuits</i>	592
Figure 154 : Répartition des mots-outils selon les <i>Nuits</i>	593
Figure 155 : Verbes : courbes des effectifs réels et théoriques.....	599
Figure 156 : Verbes : écart réduit entre effectifs réels et effectifs théoriques.....	600
Figure 157 : Noms primitifs : courbes des effectifs réels et théoriques.....	600
Figure 158 : Noms primitifs: écart réduit entre effectifs réels et effectifs théoriques.....	601
Figure 159 : Adjectifs : courbes des effectifs réels et théoriques.....	602
Figure 160 : Adjectifs : écart réduit entre effectifs réels et effectifs théoriques.....	602
Figure 161 : Noms dérivés : courbes des effectifs réels et théoriques.....	603

Figure 162 : Noms dérivés: écart réduit entre effectifs réels et effectifs théoriques.....	604
Figure 163 : Mots-outils : courbes des effectifs réels et théoriques.....	605
Figure 164 : Mots-outils: écart réduit entre effectifs réels et effectifs théoriques.....	605
Figure 165 : Préambule : Déficit et excédent en catégories lexicales.....	606
Figure 166 : Nuit 1 : Déficit et excédent en catégories lexicales.....	607
Figure 167 : Nuit 2 : Déficit et excédent en catégories lexicales.....	608
Figure 168 : Nuit 3 : Déficit et excédent en catégories lexicales.....	608
Figure 169 : Nuit 4 : Déficit et excédent en catégories lexicales.....	609
Figure 170 : Nuit 5 : Déficit et excédent en catégories lexicales.....	609
Figure 171 : Nuit 6 : Déficit et excédent en catégories lexicales.....	610
Figure 172 : Nuit 7 : Déficit et excédent en catégories lexicales.....	610
Figure 173 : Nuit 8 : Déficit et excédent en catégories lexicales.....	611
Figure 174 : Nuit 9 : Déficit et excédent en catégories lexicales.....	611
Figure 175 : Nuit 10 : Déficit et excédent en catégories lexicales.....	612
Figure 176 : Nuit 13 : Déficit et excédent en catégories lexicales.....	612
Figure 177 : Nuit 14 : Déficit et excédent en catégories lexicales.....	613
Figure 178 : Nuit 15 : Déficit et excédent en catégories lexicales.....	613
Figure 179 : Nuit 16 : Déficit et excédent en catégories lexicales.....	614
Figure 180 : Graphique de dispersion des données entre effectifs réels et effectifs théoriques : Les verbes.....	617
Figure 181 : Graphique de dispersion des données entre effectifs réels et effectifs théoriques : Les noms primitifs.....	617
Figure 182 : Graphique de dispersion des données entre effectifs réels et effectifs théoriques : Les adjectifs.....	618
Figure 183 : Graphique de dispersion des données entre effectifs réels et effectifs théoriques : Les noms dérivés.....	619
Figure 184 : Graphique de dispersion des données entre effectifs réels et effectifs théoriques : Les mots-outils.....	619
Figure 185 : Analyse factorielle des 15 Nuits d' <i>Al-Ḥimtâʿ wa-l-Muḥâna</i> par 5 catégories lexicales.....	625
Figure 186 : Connexion lexicale des vocables.....	635