

## الملاحق

ملحق (أ)

مخازن معالجة اللغة العربية ومصادر للتواصل

١. أ مخازن (REPOSITORIES)

١, ١. أ موزعو المصادر

العنوان	الاسم بالإنجليزية	الاسم بالعربية
<a href="http://www ldc.upenn.edu/">http://www ldc.upenn.edu/</a>	Linguistic Data Consortium (LDC)	اتحاد البيانات اللغوية
<a href="http://www.elda.org/">http://www.elda.org/</a>	Evaluation and Language Resource Distribution Agency (ELDA)	وكالة تقويم وترويج الذخائر اللغوية
<a href="http://www.elra.info/">http://www.elra.info/</a>	European Language Resources Association (ELRA)	رابطة مصادر اللغات الأوربية

١, ٢. أ مخزن للأوراق البحثية

العنوان	الاسم بالإنجليزية	الاسم بالعربية
<a href="http://aclweb.org/anthology-new/">http://aclweb.org/anthology-new/</a>	Anthology of the Association for Computational Linguistics (ACL Anthology)	مختارات من جمعية اللسانيات الحاسوبية
<a href="http://www.mt-archive.info/">http://www.mt-archive.info/</a>	Machine Translation Archive (MT Archive)	أرشيف الترجمة الآلية

<p><a href="http://www.lrec-conf.org/">http://www.lrec-conf.org/</a></p>	<p>Proceedings of some of the Language Resources and Evaluation Conferences (LREC)</p>	<p>بعض مداوالات مؤتمر موارد اللغة والتقييم (LREC)</p>
--	--	---

١,٣ أ. مجموعة روابط

العنوان	الاسم بالإنجليزية	الاسم بالعربية
<p><a href="http://www.globalwordnet.org/AWN/Resources.html">http://www.globalwordnet.org/AWN/Resources.html</a></p>	<p>The Arabic WordNet Project's webpage on Arabic NLP Resources</p>	<p>صفحة مشروع شبكة الكلمات العربية WordNet التي تحتوي على موارد لمعالجة اللغة العربية</p>
<p><a href="http://www.elsnet.org/arabiclist.html">http://www.elsnet.org/arabiclist.html</a></p>	<p>Elsnet's List of pointers to Arabic and other Semitic NLP and Speech sites</p>	<p>موقع Elsnet الذي يحتوي على قائمة من الروابط لمواقع لمعالجة اللغة السامية والعربية والكلام وغيرها.</p>

<a href="http://www.medar.info/The_Nemlar_Project/Publications/NEMLAR-REPORT-SURVEY-FINAL_web.pdf">http://www.medar.info/The_Nemlar_Project/Publications/NEMLAR-REPORT-SURVEY-FINAL_web.pdf</a> <a href="http://www.medar.info/Archive/">http://www.medar.info/Archive/</a>	MEDAR: 2009 Arabic HLT Survey, 2005 Arabic HLT Survey, BLARK and Archive	موقع ميدار، الذي يحتوي على مصادر بحثية مثل مسح جهود المعالجة للغة العربية عام وأرشيف BLARK.
<a href="http://www.comp.leeds.ac.uk/eric/latifa/">http://www.comp.leeds.ac.uk/eric/latifa/</a>	Latifa Al-Sulaiti's webpage on Arabic Resources	صفحة لطيفة السليطي للموارد العربية
<a href="http://universal.elra.info/">http://universal.elra.info/</a>	ELRA's Universal Catalog (search for Arabic)	كتالوج ELRA العالمي (ابحث عن العربية)
<a href="http://linguistlist.org">http://linguistlist.org</a>	Linguistlist on Arabic	القائمة اللغوية في اللغة العربية
<a href="http://www-nlp.stanford.edu/links/statnlp.html">http://www-nlp.stanford.edu/links/statnlp.html</a>	Stanford's webpage on Resources for NLP	صفحة ستانفورد لموارد معالجة اللغة الطبيعية
<a href="http://language-directory.50webs.com/languages/arabic.htm">http://language-directory.50webs.com/languages/arabic.htm</a>	Arabic Language Directory	دليل اللغة العربية
<a href="http://www1.ccls.columbia.edu/~cadi m/">http://www1.ccls.columbia.edu/~cadi m/</a>	Columbia's Arabic Dialect Modeling Group	مجموعة كولومبيا لنمذجة اللهجة للغة العربية

<a href="http://www.sakhr.com/mt.aspx">http://www.sakhr.com/mt.aspx</a>	Ajeeb – links to different sites related to Sakhr's Arabic solutions	مترجم عجيب من صخر
<a href="http://www.perseus.tufts.edu/hopper/collection?collection=Perseus:collection:Arabic">http://www.perseus.tufts.edu/hopper/collection?collection=Perseus:collection:Arabic</a>	The Perseus Project's collection of Arabic materials	مشروع بيرسيوس في جمع موارد اللغة العربية
<a href="https://sites.google.com/site/nlp4arabic/">https://sites.google.com/site/nlp4arabic/</a>	NLP-4-Arabic webpage	صفحة معالجة اللغة للعربية

## ٢. أ. التواصل والمؤتمرات (NETWORKING AND CONFERENCES)

### ١, ٢. أ. الشبكات المهنية

العنوان	الاسم بالإنجليزية	الاسم بالعربية
<a href="http://www.nemlar.org/">http://www.nemlar.org/</a>	Network for Euro-Mediterranean Language Resources (NEMLAR)	شبكة الموارد الأورو-متوسطية (نملار)
<a href="http://www.medar.info/">http://www.medar.info/</a>	Mediterranean Arabic Language and Speech Technology (MEDAR)	المشروع المتوسطي لتقنيات اللغة العربية المكتوبة والمنطوقة (ميدار)
<a href="http://www.elsnet.org/">http://www.elsnet.org/</a>	European Network of Excellence in Human Language Technologies (ELSNET)	الشبكة الأوروبية للتميز في تقنيات اللغة البشرية (ELSNET)
<a href="http://en.wikipedia.org/wiki/Academy_of_the_Arabic_Language">http://en.wikipedia.org/wiki/Academy_of_the_Arabic_Language</a>	The Arabic Language Academies	أكاديميات اللغة العربية

<a href="http://listserv.linguistlist.org/cgi-bin/wa?A0=ARABIC-L">http://listserv.linguistlist.org/cgi-bin/wa?A0=ARABIC-L</a>	ARABIC-L: Arabic Language and Linguistics Mailing List	القائمة البريدية للغة العربية وعلم اللغة
<a href="https://cs.haifa.ac.il/mailman/listinfo/semitic">https://cs.haifa.ac.il/mailman/listinfo/semitic</a>	Semitic Mailing List	القائمة البريدية للسامية
<a href="http://linguistlist.org/">http://linguistlist.org/</a>	Linguist List	قائمة اللغويين
<a href="http://gandalf.aksis.uib.n o/corpora/">http://gandalf.aksis.uib.n o/corpora/</a>	Corpora Mailing List	قائمة المدونات البريدية

## ٢،٢ أ. مؤتمرات وورش عمل

العنوان	الاسم بالإنجليزية	الاسم بالعربية
<a href="http://www.semitic.tk/">http://www.semitic.tk/</a>	Workshop on Computational Approaches to Semitic Languages (LREC 2010, EACL 2009, ACL 2007, ACL 2005, ACL 2002, ACL 1998)	ورشة الطرق الحاسوبية للغات السامية
<a href="https://sites.google.com/site/spmrl2010/">https://sites.google.com/site/spmrl2010/</a>	Workshop on Statistical Parsing of Morphologically Rich Languages (SPMRL – NAACL 2010)	ورشة التحليل الإحصائي للغات الغنية صرفياً
<a href="http://www8.georgetown.edu/college/gurt/2010/index.html">http://www8.georgetown.edu/college/gurt/2010/index.html</a>	Georgetown University Round Table on Arabic Language and Linguistics (GURT 2010)	المائدة المستديرة لجامعة جورج تاون حول اللغة العربية واللغويات
N/A- غير متوفر	Workshop on Computational Approaches to Arabic Script-based Languages (MTSummit XII 2009, LSA 2007, COLING 2004)	ورشة الطرق الحاسوبية للغات المعتمدة على الخط العربي
<a href="http://www.elda.org/medar-conference/">http://www.elda.org/medar-conference/</a>	Conference on Arabic Language Resources and Tools (MEDAR-2009, NEMLAR-2004)	مؤتمر موارد وأدوات اللغة العربية

<a href="http://www.iscal.org.sa">http://www.iscal.org.sa</a>	International Symposium on Computer and Arabic Language (ISCAL 2009, ISCAL 2007)	الندوة الدولية عن الحاسب واللغة العربية
<a href="http://www.lrec-conf.org/proceedings/lrec2008/workshops/W18_Proceedings.pdf">http://www.lrec-conf.org/proceedings/lrec2008/workshops/W18_Proceedings.pdf</a>	Workshop on HLT & NLP within the ArabWorld (LREC 2008)	ورشة معالجة اللغة الطبيعية في العالم العربي
<a href="http://infos2010.fci.cu.edu.eg/">http://infos2010.fci.cu.edu.eg/</a>	NLP track in the International Conference on Informatics and Systems (INFOS 2010, INFOS 2008)	مسار معالجة اللغة في المؤتمر الدولي للمعلوماتية والأنظمة
N/A- غير متوفر	Colloque International sur leTraitementAutomatique de la Langue Arabe (CITALA) (Rabat, 2007)	الندوة الدولية حول اللغة العربية
<a href="http://www.mt-archive.info/BCS-2006-TOC.htm">http://www.mt-archive.info/BCS-2006-TOC.htm</a>	The Challenge of Arabic for NLP/MT Conference (British Computer Society 2006)	مؤتمر تحديات معالجة اللغة العربية والترجمة الآلية
<a href="http://old-site.clsp.jhu.edu/ws2005/groups/arabic/">http://old-site.clsp.jhu.edu/ws2005/groups/arabic/</a>	Parsing Arabic Dialects (Johns Hopkins University summer workshop 2005)	تحليل اللهجات العربية
<a href="http://aune.lpl.univ-aix.fr/jep-taln04/proceed/actes/arabe.htm">http://aune.lpl.univ-aix.fr/jep-taln04/proceed/actes/arabe.htm</a>	Special session on Arabic Processing inTraitementAutomatique du LangageNaturel (TALN 2004)	عدد خاص في معالجة اللغة العربية
<a href="http://www.amtaweb.org/summit/WS2/MT_Summit_IX_Semitic_Workshop.htm">http://www.amtaweb.org/summit/WS2/MT_Summit_IX_Semitic_Workshop.htm</a>	Workshop on Machine Translation for Semitic Languages (MT Summit 2003)	ورشة عمل حول الترجمة الآلية للغات السامية

<a href="http://old-site.clsp.jhu.edu/ws2002/groups/arabic/">http://old-site.clsp.jhu.edu/ws2002/groups/arabic/</a>	Novel Approaches to Arabic Speech Recognition ( Johns Hopkins University summer workshop 2002)	طرق جديدة للتعرف على الكلام باللغة العربية
<a href="http://www.lrec-conf.org/lrec2002/lrec/wksh/ArabicLRs.html">http://www.lrec-conf.org/lrec2002/lrec/wksh/ArabicLRs.html</a>	Workshop on Arabic Language Resources and Evaluation (LREC 2002)	ورشة حول مصادر وتقييم اللغة العربية
<a href="http://www.elsnet.org/acl2001-arabic.html">http://www.elsnet.org/acl2001-arabic.html</a>	Workshop on Arabic Language Processing (ACL/EACL-2001)	ورشة حول معالجة اللغة العربية
N/A- غير متوفر	Arabic Translation and Localization Symposium (ATLAS 1999)	ندوة الترجمة العربية والتوطين

## ملحق ب: كتب ومصادر لمعالجة اللغة العربية الطبيعية

## ١. ب لغويات (المصادر باللغة الإنجليزية)

- Badawi, E., M. G. Carter, and A. Gully. 2004. Modern Written Arabic: A Comprehensive Grammar. Routledge, London.
- Bateson, Mary Catherine. 2003. Arabic Language Handbook. Georgetown University Press.
- Brustad, Kristen E. 2000. The Syntax of Spoken Arabic: A Comparative Study of Moroccan, Egyptian, Syrian, and Kuwaiti Dialects. Georgetown University Press.
- Fischer, W. 2001. A Grammar of Classical Arabic. Yale Language Series. Yale University Press. Translated by Jonathan Rodgers.
- Holes, Clive. 2004. Modern Arabic: Structures, Functions, and Varieties. Georgetown University Press.
- Georgetown Classics in Arabic Language/Linguistics. Karin C. Ryding and Margaret Nydell, eds. (several excellent titles)
- Ryding, Karin C. 2006. A Reference Grammar of Modern Standard Arabic. Cambridge University Press.
- Schulz, Eckehard. 2008. A Student Grammar of Modern Standard Arabic. Cambridge University Press.
- Buckley, Ron. 2004. Modern Literary Arabic: A Reference Grammar. Librairie du Liban.
- Wright, W. 1896. A Grammar of the Arabic Language. Cambridge University Press. (a classic grammar book)
- Bohas, G., J. Guillaume and D. Kouloughli. 2006. The Arabic Linguistic Tradition. Georgetown University Press.
- Almulla Traditional Arabic Grammar Rules : (أكثر من ٨٠٠ شريحة باللغة العربية)

## ٢. ب قواميس ورقية وممسوحة (المصادر باللغة الإنجليزية)

- قائمة بالقواميس العربية من دار العلم للملايين (قواميس أحادية وثنائية وثلاثية):

الموقع : <http://www.malayin.com/laut.asp?catid=2>

- Wehr, Hans. 1979. Dictionary of Modern Written Arabic (Arabic-English). Ithaca: Spoken Language Services.
- Lane, Edward William. 1968. Arabic-English Lexicon. Beirut: Librarie du Liban.
- Hinds, Martin and El-Said Badawi. 1986. A Dictionary of Egyptian Arabic. Beirut: Librarie du Liban.
- Stowasser, Karl and Moukhtar Ani. 2004. A Dictionary of Syrian Arabic. Georgetown University Press.
- Clarity, B. E., Karl Stowasser, Ronald G. Wolfe, D. R. Woodhead, and Wayne Beene. 2003. A Dictionary of Iraqi Arabic. Georgetown University Press.
- Harrell, Richard S. and Harvey Sobelman. 2004. A Dictionary of Moroccan Arabic. Georgetown University Press.
- Qafisheh, Hamdi. 1999. NTC's Gulf Arabic - English Dictionary. NTC Publishing Group.
- Qafisheh, Hamdi. 1999. NTC's Yemeni Arabic - English Dictionary. NTC Publishing Group.

### ٣. ب اللغويات الحاسوبية (المصادر باللغة الإنجليزية)

- Kiraz, George. 2001. Computational Nonlinear Morphology with Emphasis on Semitic Languages. Studies in Natural Language Processing. Cambridge University Press.
- Van den Bosch, A. and A. Soudi. 2007. Arabic Computational Morphology: Knowledge-based and Empirical Methods. Springer.
- Farghaly, Ali. 2010. Arabic Computational Linguistics. The University of Chicago Press.
- Farghaly, Ali and Khaled Shaalan. 2009. Arabic Natural Language Processing: Challenges and Solutions. A Special Issue of the ACM Transactions on Asian Language Information Processing (TALIP).
- Wintner, Shuly. 2009. Language Resources for Semitic Languages – Challenges and Solutions. In Sergei Nirenburg (ed.) Language Engineering for Lesser-Studied Languages. Amsterdam: IOS Press.
- United Nations Report. 2003. Harmonization of ICT standards related to Arabic Language use in information society applications. (A very informative report on information and communication technology in the Arab World).

## ٤. ب دروس ومحاضرات (المصادر باللغة الإنجليزية)

الاسم	عنوان الدرس أو المحاضرة	الرابط
نزار حبش	مقدمة في معالجة اللغة العربية الطبيعية Introduction to Arabic Natural Language Processing.	<a href="http://www1.ccls.columbia.edu/~cadiam/TUTORIAL.ARABIC.NLP.pdf">http://www1.ccls.columbia.edu/~cadiam/TUTORIAL.ARABIC.NLP.pdf</a>
Otakar Smrž	مقدمة أخرى في معالجة اللغة العربية الطبيعية Yet Another Introduction to Arabic Natural Language Processing.	<a href="http://ufal.mff.cuni.cz/~smrz/ANLP/anlp-lecture-notes.pdf">http://ufal.mff.cuni.cz/~smrz/ANLP/anlp-lecture-notes.pdf</a>
منى دياب و نزار حبش	معالجة اللهجات العربية Arabic Dialect Processing.	<a href="http://www1.ccls.columbia.edu/~cadiam/ArabicDialectTutorialAMTA2006.pdf">http://www1.ccls.columbia.edu/~cadiam/ArabicDialectTutorialAMTA2006.pdf</a>
نزار حبش	اللغات السامية الظواهر والاختلافات Semitic Linguistic Phenomena and Variations.	<a href="http://www.amtaweb.org/summit/WS2/Invited_Talk_Habash.ppt">www.amtaweb.org/summit/WS2/Invited_Talk_Habash.ppt</a>

## ملحق ج: مدونات ومعاجم لمعالجة اللغة العربية الطبيعية

## ١. ج مدونات الكلام

الملف النصي	المصدر الصوتي	المدونة
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T20">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T20</a>	<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S46">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S46</a>	أخبار الإذاعة العربية Arabic Broadcast News
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97T19">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97T19</a>	<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97S45">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC97S45</a>	الكلام العربي المصري CALLHOME Egyptian Arabic Speech
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002T39">www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002T39</a>	<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002S22">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002S22</a>	المحادثات الهاتفية المصرية العربية Egyptian Arabic telephone conversations: 1997 HUB5 Arabic Evaluation
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T15">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T15</a>	<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S43">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S43</a>	محادثات هاتفية للعربية الخليجية Appen's Gulf Arabic Conversational Telephone Speech

<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S45">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S45</a></p>	<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T16">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T16</a></p>	<p>محادثات هاتفية للعربية العراقية Appen's Iraqi Arabic Conversational Telephone Speech</p>
<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T01">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T01</a></p>	<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007S01">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007S01</a></p>	<p>محادثات هاتفية للعربية الشامية Appen's Levantine Arabic Conversational Telephone Speech</p>
<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T07">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T07</a></p>	<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S29">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S29</a></p>	<p>مزيج من بيانات التدريب لأربع مجموعات بواقع ٢٥٠ ساعة من الاتصال الهاتفي بالعربية الشامية Levantine Arabic QTTraining Data Set 5</p>
<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005S08">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005S08</a></p>		<p>BBN/AUB DARPA Babylon Levantine Arabic Speech and Transcripts</p>
<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002S02">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002S02</a></p>		<p>West Point Arabic Speech</p>
<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC96S49">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC96S49</a></p>		<p>كلم صديقاً بالعربية المصرية CALLFRIEND Egyptian Arabic</p>

<a href="http://www.mghamdi.com/SAAVB_KSU.pdf">http://www.mghamdi.com/SAAVB_KSU.pdf</a>	بنك الأصوات للهجات السعودية Saudi Accented Arabic Voice Bank (SAAVB)
<a href="http://www.mghamdi.com/KAPD.PDF">http://www.mghamdi.com/KAPD.PDF</a>	قاعدة بيانات مدينة الملك عبدالعزیز للصوتيات العربية KACST Arabic Phonetics Database (KAPD)
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة الأخبار الإذاعية العربية (الصوت والنص) NetDC Arabic Broadcast News Speech Corpus
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	GlobalPhone Arabic
<a href="http://projects ldc.upenn.edu/EARS/Arabic/">http://projects ldc.upenn.edu/EARS/Arabic/</a>	إرشادات النسخ بالعربية Arabic Transcription Guidelines in EARS

<p><a href="http://projects ldc.upenn.edu/EARS/Arabic/EARS_AMADAT.htm">http://projects ldc.upenn.edu/EARS/Arabic/EARS_AMADAT.htm</a></p>	<p>أداة النسخ العربية متعددة اللهجات AMADAT: Arabic Multi-Dialectal Transcription tool</p>
<p><a href="http://catalog.elra.info/product_info.php?products_id=874">http://catalog.elra.info/product_info.php?products_id=874</a></p>	<p>مدونة نملار للأخبار الإذاعية العربية NEMLAR Broadcast News Speech Corpus</p>
<p><a href="http://catalog.elra.info/product_info.php?products_id=875">http://catalog.elra.info/product_info.php?products_id=875</a></p>	<p>مدونة نملار لتوليف الكلام NEMLAR Speech Synthesis Corpus</p>
<p><a href="http://www.speechdat.org/ORIENTEL/index.html">http://www.speechdat.org/ORIENTEL/index.html</a></p>	<p>OrienTel هو المشروع الأوروبي الذي يركز على تطوير الموارد اللغوية للتطبيقات القائمة على الكلام</p>

## ٢. ج مدونة لتمييز الكتابة اليدوية العربية وتقييمها

العنوان	المصدر
<a href="http://www.elda.org/medar-conference/pdf/70.pdf">http://www.elda.org/medar-conference/pdf/70.pdf</a>	مصادر اتحاد البيانات اللغوية لتمييز الكتابة اليدوية العربية LDC Resources for Arabic Handwriting Recognition
<a href="http://appliedmediaanalysis.com/Datasets.htm">http://appliedmediaanalysis.com/Datasets.htm</a>	تحليل مجموعة البيانات التطبيقية لوسائل الإعلام العربية المكتوبة بخط اليد Applied Media Analysis dataset for handwritten Arabic
<a href="http://www.nist.gov/itl/iad/mig/hart2010.cfm">http://www.nist.gov/itl/iad/mig/hart2010.cfm</a>	التعرف على الكتابة اليدوية المفتوحة وتقييم الترجمة The 2010 NIST Open Handwriting Recognition and Translation Evaluation

## ٣. ج المدونات النصية

## ١، ٣. ج النص الأحادي

العنوان	المصدر
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T30">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T30</a>	Arabic Gigaword
<a href="http://www.comp.leeds.ac.uk/eric/latifa/research.htm">http://www.comp.leeds.ac.uk/eric/latifa/research.htm</a>	مدونة اللغة العربية الحديثة Corpus of Contemporary Arabic
<a href="http://arabiccorpus.byu.edu/">http://arabiccorpus.byu.edu/</a>	المدونة العربية من جامعة (BYU)

<a href="http://corpus.quran.com/">http://corpus.quran.com/</a>	المدونة العربية القرآنية (توسيم لأقسام الكلام والنحو) Quranic Arabic Corpus
<a href="http://www.almeshkat.net/books/index.php">http://www.almeshkat.net/books/index.php</a>	مكتبة المشكاة (يحتوي على مجموعة من الكتب العربية النصية المجانية)
<a href="http://www.al-eman.com/%D8%A7%D9%84%D9%83%D8%AA%D8%A8/p1">http://www.al-eman.com/%D8%A7%D9%84%D9%83%D8%AA%D8%A8/p1</a>	مجموعة نداء الإيمان (يحتوي على مجموعة من الكتب العربية المجانية)
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة الحياة العربية Al-Hayat Arabic Corpus
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة جريدة النهار النصية An-Nahar Newspaper Text Corpus
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة ليموند دبلماتك العربية Le Monde Diplomatique Text corpus in Arabic

<a href="http://www.qamus.org/corpus.htm">http://www.qamus.org/corpus.htm</a>	صفحة القاموس للمدونات
---	--------------------------

## ٢,٣. ج النص الموازي

العنوان	المصدر
الجزء الأول <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T24">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T24</a>	النص الموازي لأخبار إذاعة عربية: GALE
الجزء الثاني <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2008T09">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2008T09</a>	المرحلة ١ Arabic Broadcast News Parallel Text: GALE Phase 1
الجزء الأول <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T03">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T03</a>	النص الموازي لمجموعة الأخبار العربية: GALE
الجزء الثاني <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T09">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T09</a>	المرحلة ١ Arabic Newsgroup Parallel Text: GALE Phase 1
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2008T02">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2008T02</a>	نص المدونة العربية الموازي: GALE المرحلة ١ Arabic Blog Parallel Text: GALE Phase 1

<a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T08">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T08</a>	<p>ISI العربية الإنجليزية النص الموازي المستخرج تلقائياً (وكالة الأنباء) ISI Arabic-English Automatically Extracted Parallel Text (newswire)</p>
<a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T18">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T18</a>	<p>الأخبار الموازية بالعربية الإنجليزية : الجزء ١ Arabic English Parallel News Part 1</p>
<a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T17">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T17</a>	<p>نص ترجمة الأخبار العربية الجزء ١ Arabic News Translation Text Part 1</p>
<a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T22">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T22</a>	<p>مجموعة ترجمة وكالة الأنباء العربية (بيانات) (PATB Arabic Newswire English Translation Collection (PATB data)</p>

<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T10">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T10</a></p>	<p>البنك الشجري للإنجليزية العربية الإصدار ١,٠ (البنك الشجري من الترجمة الإنجليزية لأجزاء من بنك بنسلفانيا الشجري بالعربية ١ الجزء الإصدار ٣,٠)</p> <p>English-Arabic Treebank v 1.0 (Treebank of English translation of portions of Penn Arabic Treebank Part 1 version 3.0)</p>
<p>الجزء الأول <a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2003T18">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2003T18</a></p>	<p>العربية متعددة الترجمة (MTA)</p>
<p>الجزء الثاني <a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T05">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T05</a></p>	<p>Multiple-Translation Arabic (MTA)</p>
<p><a href="http://news meedan net/">http://news meedan net/</a></p>	<p>ميدان: منتدى مواضيع موازية للعربية - الإنجليزية</p>

documents.un.org	<p>نظام الوثائق الرسمية للأمم المتحدة (الإنجليزية، الفرنسية، الإسبانية والعربية والروسية والصينية) Official Document System of the United Nations (English, French, Spanish, Arabic, Russian, Chinese)</p>
http://ar.wikipedia.org	<p>ويكيبيديا العربية مع العديد من المصطلحات المقترنة مع لغات أخرى (غير متوازية تماما) Arabic Wikipedia with many terms paired with other languages (not strictly parallel)</p>
http://arz.wikipedia.org	<p>ويكيبيديا بالمصري Egyptian Arabic Wikipedia</p>
http://www.sacred-texts.com/isl/htq/index.htm	<p>القرآن بالعربية مع أربع ترجمات The Qur'an in Arabic with four translations</p>
http://www.sacred-texts.com/bib/wb/arb/index.htm	<p>الكتاب المقدس باللغة العربية</p>

<a href="http://www.sacred-texts.com/bib/kjv/index.htm">http://www.sacred-texts.com/bib/kjv/index.htm</a>	الكتاب المقدس باللغة الإنجليزية
<a href="http://www.sacred-texts.com/bib/wb/index.htm">http://www.sacred-texts.com/bib/wb/index.htm</a>	الكتاب المقدس بلغات أخرى
<a href="http://www.biblegateway.com/versions/?action=getVersionInfo&amp;vid=28">http://www.biblegateway.com/versions/?action=getVersionInfo&amp;vid=28</a>	الكتاب المقدس مترجم ومشكل باللغة العربية (بواسطة Biblica) A fully diacritized modern Arabic translation of the Bible (by Biblica).
<a href="http://www.umiacs.umd.edu/~resnik/strand/strand_enar.0702.txt">http://www.umiacs.umd.edu/~resnik/strand/strand_enar.0702.txt</a>	قاعدة بيانات STRAND ثنائية اللغة عربي - إنجليزي (أزواج من عناوين URL جمعت (آليا) The STRAND Arabic- English Bilingual Database (automatically collected pairs of URLs)

## ٣,٣. ج النص المعنون بأقسام الكلام أو/و المشكل

العنوان	المصدر
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة ليموند دبلوماسك العربية المعنونة (نص خام، مشكلة، معنونة بأقسام الكلام) "Le Monde Diplomatique" Arabic tagged corpus
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	مدونة نملار المكتوبة (نص خام، نص مع التحليل العربي المعجمي ، نص معنون بأقسام الكلام) NEMLAR Written Corpus: raw text, fully vowelized text, text with Arabic lexical analysis, text with Arabic POS-tags
<a href="http://zeus.cs.pacificu.edu/shereen/research.htm#corpora">http://zeus.cs.pacificu.edu/shereen/research.htm#corpora</a>	مدونة خوجة معنونة بأقسام الكلام Khoja POS tagged corpus
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T30">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T30</a>	البنك الشجري للتحليل النحوية العربية: جزء ٤ إصدار ١.٠ (فقط الصرف وسمات أقسام الكلام) Arabic Treebank: Part 4 v 1.0 (only morphology and POS tags)
<a href="http://cl.haifa.ac.il/projects/quran/index.shtml">http://cl.haifa.ac.il/projects/quran/index.shtml</a>	قرآن جامعة حيفا المعنونة مرفولوجيا University of Haifa Morphologically Tagged Qur'an

## ٤, ٣. ج التحشية لاستخراج واسترجاع المعلومات

العنوان	المصدر
<a href="http://www.itl.nist.gov/iad/mig/tests/ace/">http://www.itl.nist.gov/iad/mig/tests/ace/</a>	استخراج المحتوى
<b>(2004)</b> <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T09">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T09</a>	التلقائي : مدونة تدريب متعددة اللغات
<b>(2005)</b> <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T09">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T09</a>	Automatic Content Extraction Evaluation (ACE) Multilingual Training Corpus
<b>(2006)</b> <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T06">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006T06</a>	
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T20">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2007T20</a>	المرحلة ١ من GALE بيانات التقطير للتدريب GALE Phase 1 Distillation Training data

## ٥, ٣. ج البنوك الشجرية للتحاليل النحوية

العنوان	المصدر
<a href="http://www.ircs.upenn.edu/arabic/">http://www.ircs.upenn.edu/arabic/</a>	بنك بنسلفانيا الشجري للتحاليل النحوية Penn Arabic Treebank
<b>Part 1 v 3.0</b> <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T02">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T02</a>	بنك بنسلفانيا
<b>Part 2 v 2.0</b> <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T02">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T02</a>	الشجري للتحاليل

ogId=LDC2004T02	
<b>Part 3 v 2.0</b>	النحوية عن طريق
http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2005T20	LDC
http://www.ircs.upenn.edu/arabic/	
	بنك بنسلفانيا الشجري للتحليل النحوية الإرشادات المورفولوجية والنحوية Penn Arabic Treebank Morphological and Syntactic Guidelines
http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004T23	
	بنك براغ للتحليل
http://padt-online.blogspot.com/	الشجرية النحوية
	التبعية للغة العربية
	إصدار ١.٠
	Prague Arabic Dependency Treebank 1.0: (114K tokens) – PADT 2.0.
http://www1.ccls.columbia.edu/~CATiB/Home.html	
	كاتب: بنك
	كولومبيا للتحليل
	الشجرية النحوية
	الإصدار ١.٠
	CATiB: Columbia Arabic Treebank 1.0 (Website) (available through LDC - LDC2009E06 - by request ldc@ldc.upenn.edu)

<a href="http://www.aclweb.org/anthology/W/W09/W09-0806.pdf">http://www.aclweb.org/anthology/W/W09/W09-0806.pdf</a>	<p>جامعة مدينة دبلن بنك LFG للتحاليل الشجرية النحوية العربية Dublin City University LFG Arabic Treebank</p>
<a href="http://corpus.quran.com/treebank.jsp">http://corpus.quran.com/treebank.jsp</a>	<p>بنك جامعة ليدز للتحاليل الشجرية النحوية للقرآن University of Leeds' Quran Treebank</p>
<a href="http://papers.ldc.upenn.edu/LREC2006/PilotDialectalArabicTreebank.pdf">http://papers.ldc.upenn.edu/LREC2006/PilotDialectalArabicTreebank.pdf</a>	<p>بنك لتحاليل الشجرية النحوية للشامية العربية Levantine Arabic Treebank (available through LDC - LDC2005E78 - by request ldc@ldc.upenn.edu)</p>
<a href="http://www.lrec-conf.org/proceedings/lrec2008/pdf/880_paper.pdf">http://www.lrec-conf.org/proceedings/lrec2008/pdf/880_paper.pdf</a>	<p>بنك الأبنية الحملية العربية Arabic Proposition Bank (Propbank)</p>

<p><a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T24">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T24</a></p>	<p>الإصدار ٣,٠ من</p>
<p><b>(BBN's webpage)</b> <a href="http://www.bbn.com/ontonotes/">http://www.bbn.com/ontonotes/</a></p>	<p>OntoNotes (نصوص إنجليزية وعربية وصينية محشاة، بناء جملة المسند، ومعنى الكلمة، والإحالة) OntoNotes Release 3.0 (English, Arabic and Chinese texts annotated for syntax, predicate argument structure, word sense and coreference).</p>
<p><a href="http://beta.visl.sdu.dk/visl/ar/">http://beta.visl.sdu.dk/visl/ar/</a></p>	<p>موقع جامعة Syddansk لتعلم بناء الجملة بشكل تفاعلي مرئي Syddansk Universiteit's Visual Interactive Syntax Learning Site</p>

٤. ج مدونات التقييم

العنوان	المصدر
<p><a href="http://www.itl.nist.gov/iad/mig/tests/lre/">http://www.itl.nist.gov/iad/mig/tests/lre/</a></p>	<p>NIST تقييم التمييز اللغوي</p>
<p>بيانات LRE-2003 <a href="http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S31">http://www ldc upenn edu/Catalog/CatalogEntry.jsp?catalogId=LDC2006S31</a></p>	<p>NIST's Language Recognition Evaluation</p>

<p><b>مجموعة اختبار LRE-2007</b></p> <p><a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009S04">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009S04</a></p>	(LRE)
<p><b>التدريب التكميلي</b></p> <p><a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009S05">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009S05</a></p>	
<p><a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T05">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2009T05</a></p>	<p>قياسات NIST للترجمة الآلية (MetricsMATR08)</p> <p>بيانات التكوير 2008 NIST Metrics for Machine Translation (MetricsMATR08) Development Data</p>
<p><a href="http://www.itl.nist.gov/iad/mig/tests/mt/">http://www.itl.nist.gov/iad/mig/tests/mt/</a></p>	<p>تقييم NIST للترجمة الآلية NIST's Machine Translation Evaluation</p>
<p><a href="http://www.technolanguge.net/article.php3?id_article=199">http://www.technolanguge.net/article.php3?id_article=199</a></p>	<p>حزمة تقييم CESTA مبادرة فرنسية للترجمة الآلية عربي - فرنسي. CESTA Evaluation Package: (French initiative; Arabic-French MT)</p>

<a href="http://www.technolangu.net/article.php3?id_article=201">http://www.technolangu.net/article.php3?id_article=201</a>	<p>حزمة تقييم ARCADE II - مبادرة فرنسية للترجمة الآلية عربي - فرنسي.</p> <p>ARCADE II Evaluation Package: (French initiative; Arabic-French MT)</p>
<a href="http://www.nist.gov/itl/iad/mig/hart2010.cfm">http://www.nist.gov/itl/iad/mig/hart2010.cfm</a>	<p>مشروع OpenHaRT 2010 NIST المفتوح للتعرف على الكتابة اليدوية وتقييم الترجمة</p> <p>OpenHaRT 2010: NIST Open Handwriting Recognition and Translation Evaluation</p>

### ٥. ج قواعد بيانات معجمية

#### ٥.١ ج القواميس الأحادية

• الباحث العربي: يقدم الموقع خدمة البحث في أهم القواميس والمراجع اللغوية

العربية مثل لسان العرب والقاموس المحيط، العنوان:

<http://www.baheth.info/index.jsp>

## ٢, ٥. ج القواميس المتعددة

العنوان	المصدر
<a href="http://unterm.un.org/">http://unterm.un.org/</a>	UNTERM قاعدة بيانات الأمم المتحدة للمصطلحات UNTERM United Nations Terminology Database
<a href="http://termweb.unesco.org/">http://termweb.unesco.org/</a>	قاعدة بيانات اليونسكو للمصطلحات UNESCO Term database
<a href="http://www.emro.who.int/Unified-Medical-Dictionary.html">http://www.emro.who.int/Unified-Medical-Dictionary.html</a>	المعجم الطبي الموحد لمنظمة الصحة العالمية Unified Medical Dictionary of the World Health Organization
<a href="https://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf">https://lib-thesaurus.un.org/LIB/DHLUNBISThesaurus.nsf</a>	مكنز نظام المعلومات الببليوغرافية للأمم المتحدة UN Bibliographical Information System Thesaurus
<a href="http://www.websters-online-dictionary.org/">http://www.websters-online-dictionary.org/</a>	قاموس ويبستر متعدد اللغات Webster's Online Dictionary (multilingual)

<p><a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a></p>	<p>قاعدة بيانات قاموس الرياضة متعدد اللغات (انجليزي فرنسي يوناني العربية الألمانية البرتغالية الإسبانية) Multilingual Dictionary of Sports (English-French- Greek-Arabic-German- Spanish- Portuguese) database</p>
<p><a href="http://dictionary.sakhr.com/">http://dictionary.sakhr.com/</a></p>	<p>قاموس صخر إنجليزي عربي (عربي ⇔ الإنجليزية / الفرنسية / التركية / الألمانية) Sakhr's English-Arabic dictionary Arabic⇔English/French/T urkish/German)</p>
<p><a href="http://www.ectaco.co.uk/English-Arabic-Dictionary/">http://www.ectaco.co.uk/English-Arabic-Dictionary/</a></p>	<p>ECTACO قاموس إنجليزي عربي Ectaco English-Arabic Dictionary</p>
<p><a href="http://www.welokee.nl/welokee/arabic/">http://www.welokee.nl/welokee/arabic/</a></p>	<p>قاموس Effel عربي إنجليزي Effel Arabic-English Dictionary</p>

<a href="http://qamoose.arabeyes.org/">http://qamoose.arabeyes.org/</a>	<p>عرب آيز قاموس عربي إنجليزي Arabeyes Arabic-English QaMoose Dictionary</p>
<a href="http://www.almisbar.com/index.html">http://www.almisbar.com/index.html</a>	<p>المسبار قاموس عربي إنجليزي Almisbar Arabic-English dictionary</p>
<a href="http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:2002.02.0005">http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:2002.02.0005</a>	<p>قاموس سلموني المتقدم للمتعلم عربي إنجليزي Salmoné's Advanced Learner's Arabic-English Dictionary</p>
<a href="http://www.let.kun.nl/wba/">www.let.kun.nl/wba/</a>	<p>مشروع قاموس نيميغن هولندي عربي The Nijmegen Dutch- Arabic Dictionary Project</p>
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	<p>قاموس ثنائية فرنسي عربي وعربي فرنسي (DixAF) Bilingual Dictionary French Arabic, Arabic French (DixAF) – ELRA Catalog M0040</p>
<a href="http://arabdictionary.huji.ac.il/Matrix.Arabdictionary/MainPage.aspx">http://arabdictionary.huji.ac.il/Matrix.Arabdictionary/MainPage.aspx</a>	<p>قاموس Ayelon-Shin'ar عربي عبري Ayelon-Shin'ar Arabic- Hebrew online dictionary</p>

<a href="http://www.lexicool.com/">http://www.lexicool.com/</a>	<p>Lexicool دليل قواميس ومعاجم - مقالات عديدة للعربية.</p> <p>Lexicool: A directory of online dictionaries and glossaries - numerous Arabic entries</p>
---	---

## ٣, ٥. ج المعاجم المورفولوجية

العنوان	المصدر
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC99L22">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC99L22</a>	<p>معجم العامية المصرية العربية Egyptian Colloquial Arabic Lexicon</p>
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	<p>معجم DIINAR العربي أحادي اللغة DIINAR (DIctionnaire INformatisé de l'ARabe) Monolingual Arabic Lexicon - ELRA Catalog L0073 - 120,000 lemmas. [204]</p>
انظر أيضاً المعاجم من المحللات الصرفية	

## ٤, ٥. ج قوائم الجذر

العنوان	المصدر
<a href="http://www.angelfire.com/tx4/lisan/roots1.htm">http://www.angelfire.com/tx4/lisan/roots1.htm</a>	<p>قائمة باكوالتر للجذور العربية Buckwalter's list of Arabic roots</p>
<a href="http://www.studyquran.co.uk/PRLonline.htm">http://www.studyquran.co.uk/PRLonline.htm</a>	<p>مشروع قائمة الجذور Project Root List</p>

<a href="http://tides.umiacs.umd.edu/software.html">http://tides.umiacs.umd.edu/software.html</a>	<p>قائمة الجذر داخل محلل سبحاوي الصرفي (الاتصال بالدكتور كريم درويش) Root list inside the morphological analyzer Sebawai (Contact Dr. Kareem Darwish)</p>
---	---

### ٥,٥. ج قواعد البيانات الصوتية

العنوان	المصدر
<a href="http://catalog.elra.info/search.php">http://catalog.elra.info/search.php</a>	<p>معجم LC-STAR الصوتية القياسية العربية LC-STAR Standard Arabic Phonetic Lexicon ELRA Catalog S0247</p>
<a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC99L22">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC99L22</a>	<p>معجم العامية المصرية العربية (يحتوي على بيانات صوتية وفونيمية) Egyptian Colloquial Arabic Lexicon (contains phonetic and phonemic representations)</p>

## ٥,٦. ج الفهارس الجغرافية

العنوان	المصدر
<a href="http://www.fao.org/termportal/en/">http://www.fao.org/termportal/en/</a>	FAOTERM : منظمة الأمم المتحدة للأغذية والزراعة للمصطلحات المرجعية لأسماء البلدان (بست لغات منها اللغة العربية) FAOTERM: United Nations' Food and Agriculture Organization of the Terminology reference for country names (six languages including Arabic)
<a href="http://www.foreignword.com/countries/">http://www.foreignword.com/countries/</a>	موقع Foreignword.com أسماء البلدان بـ ١٦ لغة بما فيها العربية Foreignword.com's country names in 16 languages including Arabic
<a href="http://www.geonames.de/">http://www.geonames.de/</a>	موقع Geonames.de لموارد متعددة اللغات لأسماء الكيانات الجغرافية (وأشياء أخرى) Geonames.de's multilingual resource for names of geographical entities (and other things)
<a href="http://geonames.usgs.gov/">http://geonames.usgs.gov/</a>	مجلس الولايات المتحدة المعني بالأسماء الجغرافية (بما في ذلك البلدان العربية (يستخدم SATTs للنقل الكتابي للعربية U.S. Board on Geographic Names (including Arab countries) – uses SATTs Arabic transliteration

<a href="http://www1.ccls.columbia.edu/~ybenajiba/downloads.html">http://www1.ccls.columbia.edu/~ybenajiba/downloads.html</a>	<p>ANERGazet : مجموعة من ٣ معاجم لأسماء الكيانات الجغرافية، والناس والمنظمات</p> <p>ANERGazet: a collection of 3 Gazetteers for names of geographical entities, people and organizations</p>
<a href="http://www.kanji.org/cjk/arabic/arabsam.htm">http://www.kanji.org/cjk/arabic/arabsam.htm</a>	<p>قواعد بيانات لأسماء باللغة العربية في معهد قاموس CJK</p> <p>Databases for names in Arabic at the CJK Dictionary Institute</p>
<p>LDC10001 Arabic Names (through LDC by اسم عربي عن طريق ١٠٠٠١ (request LDC2005G02</p>	

### ٥,٧ ج أنطولوجيات دلالية

- وردنت العربي، العنوان: [/http://www.globalwordnet.org/AWN](http://www.globalwordnet.org/AWN)

### ملحق د: أدوات معالجة اللغة العربية الطبيعية

#### ١. د التجذيع

العنوان	المصدر
<a href="http://zeus.cs.pacificu.edu/shereen/research.htm#stemming">http://zeus.cs.pacificu.edu/shereen/research.htm#stemming</a>	<p>أداة خوجة للتجذيع Khoja Arabic Stemmer</p>
<a href="http://tides.umiacs.umd.edu/software.html">http://tides.umiacs.umd.edu/software.html</a>	<p>المجذعات العربية: سبعاوي و مجذع إل التعريف (تواصل مع الدكتور كريم درويش)</p> <p>Arabic Stemmers: Sebawai and Al-Stem (Contact Dr. Kareem Darwish)</p>

<p><a href="http://link.springer.com/chapter/10.1007%2F978-1-4020-6046-5_12">http://link.springer.com/chapter/10.1007%2F978-1-4020-6046-5_12</a></p>	<p>مجذع اللام Larkey (تواصل مع المؤلفين) Larkey's L-stem (contact authors)</p>
--	--

٢.٥ التحليل والتوليد الصرفي

العنوان	المصدر
<p>الإصدار ١.٠ <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002L49">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2002L49</a></p>	<p>محلل باكوالتر الصرفي (باما)</p>
<p>الإصدار ١.٢ <a href="http://sourceforge.net/projects/aramorph/">http://sourceforge.net/projects/aramorph/</a></p>	<p>Buckwalter Arabic Morphological Analyzer (BAMA) versions 1.0, 1.2 and 2.0</p>
<p>الإصدار ٢.٠ <a href="http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004L02">http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC2004L02</a></p>	
<p><a href="http://www ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2010L01">http://www ldc.upenn.edu/Catalog/catalogEntry.jsp?catalogId=LDC2010L01</a></p>	<p>ساما: المحلل الصرفي العربي القياسي - الإصدار ٣ Standard Arabic Morphological Analyzer (SAMA, version 3.0 of BAMA)</p>
<p><a href="http://quest.ms.mff.cuni.cz/cgi-bin/elixir/index.fcgi">quest.ms.mff.cuni.cz/cgi-bin/elixir/index.fcgi</a></p>	<p>إلكسير إف إم (ELIXIRFM): نظام</p>
<p><a href="http://sourceforge.net/projects/elixir-fm/">http://sourceforge.net/projects/elixir-fm/</a></p>	<p>إلكسير (ELIXIR) العربي للصراف الوظيفي ElixirFM: Functional Arabic Morphology online interface (download)</p>

<a href="http://open.xerox.com/Services/arabic-morphology">http://open.xerox.com/Services/arabic-morphology</a>	<p>زيروكس للتحليل الصرفي والتوليد للغة العربية</p> <p>Xerox Arabic Morphological Analysis and Generation</p>
<a href="http://www1.ccls.columbia.edu/~habash/software-downloads/MAGEAD_Distribution.html">http://www1.ccls.columbia.edu/~habash/software-downloads/MAGEAD_Distribution.html</a>	<p>مجيد : المحلل والمولد الصرفي للعربية ولهجاتها</p> <p>MAGEAD: Morphological Analysis and Generation for Arabic and its Dialects</p>
<a href="http://www1.ccls.columbia.edu/MADA/">http://www1.ccls.columbia.edu/MADA/</a>	<p>المرجانة : المحلل والمولد الصرفي العربي المعتمد على المُعَيِّمَة</p> <p>Almorgeana: Arabic Lexeme-based Morphological Generation and Analysis is distributed as part of the MADA system.</p>
<a href="http://www.econtent.org.sa/MediaCenter/MediaCenter_DownloadsFiles/Alkhalil.rar">www.econtent.org.sa/MediaCenter/MediaCenter_DownloadsFiles/Alkhalil.rar</a>	<p>محلل الخليل الصرفي</p> <p>Alkhalil Morphological Analyzer</p>

## ٣.د فك اللبس الصرفي و تعيين أقسام الكلام

العنوان	المصدر
<a href="http://zeus.cs.pacificu.edu/shereen/research.htm#tagging">http://zeus.cs.pacificu.edu/shereen/research.htm#tagging</a>	معنون خوجة العربي Khoja Arabic Tagger
<a href="http://nlp.ldeo.columbia.edu/amira/">http://nlp.ldeo.columbia.edu/amira/</a>	أميرة: مجموعة أدوات لتقطيع النص وتعيين أقسام الكلام وتجزئة العبارات الأساسية AMIRA: Toolkit for Arabic tokenization, POS tagging and base phrase chunking
<a href="http://www1.ccls.columbia.edu/MADA/">http://www1.ccls.columbia.edu/MADA/</a>	مدى: المحلل الصرفي ومزيل الغموض العربي – أداة لتقطيع النص، تعيين المادة المعجمية، التشكيل وتعيين أقسام الكلام MADA: Morphological Analysis and Disambiguation for Arabic – a tool for tokenization, lemmatization, diacritization and POS tagging

## ٤. د المحللات

العنوان	المصدر
<a href="http://nlp.stanford.edu/software/lex-parser.shtml">http://nlp.stanford.edu/software/lex-parser.shtml</a>	محلل ستانفورد The Stanford Parser
<a href="http://www.cis.upenn.edu/~dbikel/software.html">http://www.cis.upenn.edu/~dbikel/software.html</a>	محلل بايكل The Bikel Parser
<a href="http://www.maltparser.org/">http://www.maltparser.org/</a>	محلل MALTParser
<a href="http://iness.uib.no/iness/xle-web">http://iness.uib.no/iness/xle-web</a>	محلل محمد عطية القائم على القواعد للغة العربية المعاصرة Mohammed Attia's Rule-based Parser for MSA

## ٥. د المحرر

- عربتك (ArabTeX): حزمة برمجيات حرة لدعم الألفباء العربية في تك ولاتك، العنوان:

<http://en.wikipedia.org/wiki/ArabTeX>

## ٦. د التعرف على أسماء الأعلام

العنوان	المصدر
<a href="http://www1.ccls.columbia.edu/~ybenajiba/downloads.html">http://www1.ccls.columbia.edu/~ybenajiba/downloads.html</a>	نظام ياسين بن عجيبة للتعرف على أسماء الأعلام Yassine Benajiba's ANER (Arabic Named Entity Recognition) system
<a href="http://www.bbn.com/technology/speech/identifinder">http://www.bbn.com/technology/speech/identifinder</a>	نظام BBN (إنجليزي ، عربي ، صيني) BBN's Identifinder (English, Arabic, Chinese)

## ٧. د التعديل الشجري

- برنامج (TrEd for Arabic) محرر مجاني مع دعم للغة العربية ، العنوان :  
[http://ufal.mff.cuni.cz/~pajas/arabic\\_install.html](http://ufal.mff.cuni.cz/~pajas/arabic_install.html)

## ٨. د الصناعة المعجمية

- برنامج (aConCorde) لتوليد التوافق للغة العربية ، العنوان :  
<http://www.andy-roberts.net/coding/aconcorde>

## ٩. د إدخال النص

العنوان	المصدر
<a href="http://www.yamli.com/ar/">http://www.yamli.com/ar/</a>	يملي
<a href="http://www.google.com/intl/ar/inputtools/cloud/try/">http://www.google.com/intl/ar/inputtools/cloud/try/</a>	تعريب - قوقل
<a href="http://afkar.microsoft.com/en/maren/">http://afkar.microsoft.com/en/maren/</a>	مرن - مايكروسوفت

## ١٠. د الترجمة الآلية

العنوان	المصدر
<a href="http://translate.google.com/">http://translate.google.com/</a>	ترجمة قوقل - مترجم ثنائي الاتجاه لأكثر من ٥٠ لغة منها اللغة العربية Google Translate – bidirectional translation for over 50 languages including Arabic
<a href="http://www.bing.com/translator/">http://www.bing.com/translator/</a>	مترجم مايكروسوفت - مترجم ثنائي الاتجاه لأكثر من ٣٠ لغة منها اللغة العربية Microsoft's Bing Translator – bidirectional translation for over 30 languages including Arabic
<a href="http://translate.sakhr.com">http://translate.sakhr.com</a>	ترجم من صخر (عربي - إنجليزي وإنجليزي - عربي) Sakhr's Tarjim (Arabic-English and English-Arabic)
<a href="http://www.almisbar.com/index.html">http://www.almisbar.com/index.html</a>	ترجمة المسبار عربي - إنجليزي Almisbar Arabic-English translation
<a href="http://www.statmt.org/">http://www.statmt.org/</a>	موارد عامة للترجمة الإحصائية الآلية: محاذاة الجيزة، نظام فرعون وموسى لفك التشفير، إلخ. Statistical MT public resources: Giza alignment, Pharaoh and Moses decoders, etc.

## ملحق هـ: الاختصارات الهامة في معالجة اللغة العربية الطبيعية

الإنجليزية	العربية	رقم
ACE: Automatic Content Extraction	استخراج المحتوى التلقائي	١
ACL: Association for Computational Linguistics	جمعية اللغويات الحاسوبية	٢
ACM: Association for Computing Machinery	رابطة مكائن الحوسبة	٣
AMTA: Association for Machine Translation in the Americas	جمعية الترجمة الآلية في الأمريكيتين	٤
ANLP: Applied Natural Language Processing	معالجة اللغات الطبيعية التطبيقية	٥
BC: Broadcast Conversations	المحادثات الإذاعية	٦
BLARK: Basic Language Resource Kit – a minimal set of language resources necessary to do research.	مجموعة موارد اللغة الأساسية "مجموعة صغيرة من الموارد اللغوية الضرورية للبحوث"	٧
BN: Broadcast News	أخبار الإذاعة	٨
CLIR: Cross-Language Information Retrieval	استرجاع المعلومات عبر اللغات	٩
COLABA: Cross-lingual Arabic Blogging Alerts (TSWG funded)	تنبيهات المدونات العربية عبر اللغات (تمويل TSWG)	١٠
COLING: Computational Linguistics Conference	مؤتمر اللغويات الحاسوبية	١١
COTS: Commercial-Off-The-Shelf systems (e.g., commercial Machine Translation systems)	الأنظمة التجارية الجاهزة (مثل أنظمة الترجمة الآلية التجارية)	١٢
CTS: Conversational Telephone Speech	نص المحادثات الهاتفية	١٣

DARPA: Defense Advanced Research Projects Agency (USA)	وكالة مشاريع البحوث الدفاعية المتطورة	١٤
EACL: European ACL	جمعية اللغويات الحاسوبية الأوروبية	١٥
EAMT: European Association for Machine Translation	الجمعية الأوروبية للترجمة الآلية	١٦
EARS: Efficient, Affordable, Reusable Speech-to-Text (DARPA program)	فعال- ميسر - قابل لإعادة الكلام إلى نص (برنامج وكالة مشاريع البحوث الدفاعية المتطورة)	١٧
ELDA: Evaluations and Language resources Distribution Agency	وكالة توزيع الموارد اللغوية والتقييمات	١٨
ELRA: European Language Resources Association	جمعية الموارد اللغوية الأوروبية	١٩
ELSNET: European Network of Excellence in Human Language Technologies	الشبكة الأوروبية للتميز في تقنيات اللغة البشرية	٢٠
EMNLP: Empirical Methods to Natural Language Processing Conference	مؤتمر الأساليب التجريبية لمعالجة اللغة الطبيعية	٢١
EuRADic: European and Arabic Dictionaries and Corpora	المجاميع والقواميس العربية والأوروبية	٢٢
GALE: Global Autonomous Linguistic Exploitation (DARPA program)	الاستثمار اللغوي العالمي المستقل (برنامج وكالة مشاريع البحوث الدفاعية المتطورة)	٢٣
HARD: High Accuracy Retrieval from Documents	الاسترجاع عالي الدقة من المستندات	٢٤
HDL: High-Density Language (i.e., with rich NLP resources)	اللغة عالية الكثافة ( أي مع الموارد الغنية لمعالجة اللغات الطبيعية)	٢٥
HLT: Human Language Technology	تقنية اللغة البشرية	٢٦

ICASSP: International Conference on Acoustics, Speech and Signal Processing	المؤتمر الدولي للصوتيات والكلام ومعالجة الإشارة	٢٧
ICSLP: International Conference on Spoken Language Processing	المؤتمر الدولي لمعالجة اللغة المنطوقة	٢٨
LCTL: Less Commonly Taught Languages	اللغات الأقل شيوعاً في التدريس	٢٩
LDC: Linguistic Data Consortium, University of Pennsylvania	اتحاد البيانات اللغوية (أو ائتلاف البيانات اللسانية)، جامعة بنسلفانيا	٣٠
LDL: Low-Density Language (i.e., with poor NLP resources)	اللغة منخفضة الكثافة (أي مع الموارد القليلة لمعالجة اللغات الطبيعية)	٣١
LRE: Language Recognition Evaluation	تقييم التمييز اللغوي	٣٢
LREC: Language Resources and Evaluation Conference	مؤتمر الموارد اللغوية والتقييم	٣٣
LSA: Linguistic Society of America	الجمعية اللغوية الأمريكية	٣٤
MADCAT: Multilingual Automatic Document Classification Analysis and Translation (DARPA program)	ترجمة وتحليل تصنيف المستندات التلقائية متعددة اللغات (برنامج وكالة مشاريع البحوث الدفاعية المتطورة)	٣٥
MCA: Modern Colloquial Arabic	العربية العامية الحديثة	٣٦
MEDAR: Mediterranean Arabic Language and Speech Technology, a consortium consisting of 15 partners from European and Arabic - Mediterranean countries	مدار: اللغة العربية في البحر المتوسط وتقنية الكلام، تحالف مكون من ١٥ شريكاً من الدول الأوروبية والعربية في بلدان البحر المتوسط	٣٧
MSA: Modern Standard Arabic	العربية الفصحى الحديثة	٣٨

MTS: Machine Translation Summit	قمة الترجمة الآلية	٣٩
NAACL: North American ACL	جمعية اللغويات الحاسوبية في شمال أمريكا	٤٠
NACAL: North America Conference on Afro-asiatic Languages	مؤتمر أمريكا الشمالية للغات الأفرو-آسيوية	٤١
NAPLUS: Natural Arabic Processing for Language Understanding Systems (European funded Project)	معالجة اللغة العربية الطبيعية لأنظمة فهم اللغة (مشروع بتمويل أوروبي)	٤٢
NEMLAR: Network for Euro-Mediterranean LAnguage Resources	نملار: شبكة موارد اللغة الأوروبية في البحر المتوسط	٤٣
NIST: National Institute of Standards and Technology (USA)	المعهد الوطني للمعايير والتقنية (الولايات المتحدة الأمريكية)	٤٤
NSF: National Science Foundation (USA)	المؤسسة الوطنية للعلوم (الولايات المتحدة الأمريكية)	٤٥
NW: Newswire	وكالة الأنباء	٤٦
RANLP: Recent Advances in Natural Language Processing	التطورات الأخيرة في معالجة اللغات الطبيعية	٤٧
STT: Speech-to-Text	كلام إلى نص	٤٨
SATTS: Standard Arabic Technical Transliteration System	نظام النقل الكتابي المعياري التقني للعربية الفصحى	٤٩
TALN: Traitement Automatique du Langage Naturel	معالجة اللغات الطبيعية	٥٠
TDT: Topic Detection and Tracking evaluation	كشف الموضوع وتقييم المتابعة	٥١

TIDES: Translingual Information Detection Extraction and Summarization (DARPA program)	تلخيص واستخراج وكشف المعلومات عبر اللغات (برنامج وكالة مشاريع البحوث الدفاعية المتطورة)	٥٢
TRANSTAC: Translation for Tactical Use (DARPA program)	الترجمة للاستخدام المنهجي (التكتيكي) (برنامج وكالة مشاريع البحوث الدفاعية المتطورة)	٥٣
TREC: Text Retrieval Conference	مؤتمر استرجاع النص	٥٤
TSWG: Technical Support Working Group (a US interagency research and development program)	فريق عمل الدعم الفني (بحث مشترك بين الوكالات الأمريكية وبرنامج التنمية)	٥٥
TTS: Text-to-Speech	تحويل النص لكلام	٥٦
WB: Weblogs	مدونات	٥٧

## المراجع

### References

- [1] El-Said M. Badawi. *Mustawayat al-'Arabiyya al-mu'asira fi Misr (The Levels of Modern Arabic in Egypt)*. Cairo: Dar al-Ma'arif, 1973. 1, 2
- [2] Reem Bassiouney. *Arabic Sociolinguistics: Topics in Diglossia, Gender, Identity, and Politics*. Georgetown University Press, 2009. 2
- [3] Charles F Ferguson. Diglossia. *Word*, 15(2):325–340, 1959. 2
- [4] Nizar Habash, Abdelhadi Soudi, and Tim Buckwalter. On Arabic Transliteration. In A. van den Bosch and A. Soudi, editors, *Arabic Computational Morphology: Knowledge-based and Empirical Methods*. Springer, 2007. 4, 21, 27, 31
- [5] TimBuckwalter. Issues in Arabic Morphological Analysis. In A. van den Bosch and A. Soudi, editors, *Arabic Computational Morphology: Knowledge-based and Empirical Methods*. Springer, 2007. 7, 13, 33
- [6] Kam-Fai Wong, Wenji Li, Ruifeng Xu, and Zheng sheng Zhang. *Introduction to Chinese Natural Language Processing*. Synthesis Lectures on Human Language Technologies. Morgan and Claypool, 2010. 8
- [7] Elsaid Badawi, Mike G. Carter, and Adrian Gully. *Modern Written Arabic: A Comprehensive Grammar*. Routledge, 2004. 10, 63, 93
- [8] Nizar Habash and Owen Rambow. Morphophonemic and Orthographic Rules in a Multi-Dialectal Morphological Analyzer and Generator for Arabic Verbs. In *International Symposium on Computer and Arabic Language (ISCAL)*, Riyadh, Saudi Arabia, 2007. 10, 32, 59
- [9] Mohamed Maamouri, Ann Bies, Tim Buckwalter, and Wigdan Mekki. *The Penn Arabic Treebank: Building a Large-Scale Annotated Arabic Corpus*, 2004. 11, 53, 79, 93, 104, 105
- [10] KaisDukes andNizar Habash. Morphological Annotation of Quranic Arabic. In *Proceedings of the Language Resources and Evaluation Conference (LREC)*, Malta, 2010. 12, 112
- [11] Rani Nelken and Stuart Shieber. Arabic Diacritization Using Weighted Finite-State Transducers. In *Proceedings of the Workshop on Computational Approaches to Semitic Languages at 43<sup>rd</sup> Meeting of the Association for*

- Computational Linguistics (ACL'05)*, pages 79–86, Ann Arbor, Michigan, 2005. DOI: 10.3115/1621787.1621802, 13, 24
- [12] Imed Zitouni, Jeffrey S. Sorensen, and Ruhi Sarikaya. Maximum Entropy Based Restoration of Arabic Diacritics. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 577–584, Sydney, Australia, 2006. DOI: 10.3115/1220175.1220248, 13, 24
- [13] Ahmed El Kholy and Nizar Habash. Techniques for Arabic Morphological Detokenization and Orthographic Denormalization. In *Workshop on Language Resources and Human Language Technology for Semitic Languages in the Language Resources and Evaluation Conference (LREC)*, Valletta, Malta, 2010. 13, 36, 69, 77, 78, 122, 124
- [14] Nizar Habash and Ryan Roth. Identification of Naturally Occurring Numerical Expressions in Arabic. In *Proceedings of the Language Resources and Evaluation Conference (LREC)*, Marrakech, Morocco, 2008. 14, 118
- [15] Alan Kaye. Adaptations of Arabic Script. In P.T. Daniels and W. Bright, editors, *The World's Writing Systems*. Oxford University Press, 1996. 14
- [16] Klaus Lagally. ArabTEX: Typesetting Arabic and Hebrew, User Manual Version 4.00. Technical Report 2004/03, Fakultät Informatik, Universität Stuttgart, March 11 2004. 16, 76
- [17] Nizar Habash. Nuun: A System for Developing Platform and Browser Independent Arabic Web Applications. In *Proceedings of the Arabic Translation and Localization Conference (ATLAS-99)*, Tunis, Tunisia, 1999. 17
- [18] United Nations Report. *HARMONIZATION OF ICT STANDARDS RELATED TO ARABIC LANGUAGE USE IN INFORMATION SOCIETY APPLICATIONS*. United Nations Publication, <http://unpan1.un.org/intradoc/groups/public/documents/unescwa/unpan030581.pdf>, 2003. 18
- [19] Gina Engström. Internationalisation and Localisation Problems in the Chinese and Arabic Scripts. Master's thesis, Uppsala University, 2008. 18, 19
- [20] Kenneth R. Beesley. Romanization, Transcription and Transliteration, 1997. <http://www.xrce.xerox.com/Research-Development/Historical-projects/Linguistic-Demos/Arabic-Morphological-Analysis-and-Generation/Romanization-Transcription-and-Transliteration>. 20
- [21] Y. Al-Onaizan and K. Knight. Machine Transliteration of Names in Arabic Text. In *Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages*, 2002. DOI: 10.3115/1118637.1118642, 21, 36, 120
- [22] Ulf Hermjakob, Kevin Knight, and Hal Daumé III. Name Translation in Statistical Machine Translation - Learning When to Transliterate. In *Proceedings of ACL-08: HLT*, Columbus, Ohio, 2008. 21, 36, 120
- [23] Tim Buckwalter. Buckwalter Arabic Morphological Analyzer Version 2.0, 2004. *Linguistic Data Consortium*, University of Pennsylvania. LDC Catalog No.: LDC2004L02, ISBN 1-58563-324-0. 20, 32, 36, 41, 47, 67, 68, 69, 70, 71

- [24] Fadi Biadsy, Jihad El-Sana, and Nizar Habash. Online Arabic handwriting recognition using Hidden Markov Models. In *The 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR'10)*, La Baule, France, 2006. 23
- [25] Volker Märgner and Haikal El Abed. Arabic Word and Text Recognition - Current Developments. In Khalid Choukri and Bente Maegaard, editors, In *Proceedings of the Second International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, April 2009. The MEDAR Consortium. 23
- [26] Liana M. Lorigo and Venu Govindaraju. Offline Arabic Handwriting Recognition: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5):712–724, 2006. DOI: 10.1109/TPAMI.2006.102, 23
- [27] Kareem Darwish and Douglas W. Oard. Term Selection for Searching Printed Arabic. In *SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 261–268, New York, NY, USA, 2002. ACM. DOI: 10.1145/564376.564423, 23
- [28] Walid Magdy and Kareem Darwish. Arabic OCR Error Correction Using Character Segment Correction, Language Modeling, and Shallow Morphology. In *Proceedings of 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, pages 408–414, Sydney, Australia, 2006. DOI: 10.3115/1610075.1610132, 23
- [29] Prem Natarajan, Shirin Saleem, Rohit Prasad, Ehry MacRostie, and Krishna Subramanian. Arabic and Chinese Handwriting Recognition, volume 4768 of *Lecture Notes in Computer Science*, pages 231–250. Springer, Berlin, Germany, 2008. 23
- [30] Shirin Saleem, Huaigu Cao, Krish na Subramanian, Marin Kamali, Rohit Prasad, and Prem Natarajan. Improvements in BBN's HMM-based Offline Handwriting Recognition System. In Khalid Choukri and Bente Maegaard, editors, *10th International Conference on Document Analysis and Recognition (ICDAR)*, Barcelona, Spain, July 2009. 23
- [31] Zhidong Lu, Issam Bazzi, Andras Kornai, John Makhoul, Premkumar Natarajan, and Richard Schwartz. A Robust, Language-Independent OCR System. In *The 27th AIPR Workshop: Advances in Computer Assisted Recognition*, SPIE, 1999. DOI: 10.1117/12.339811, 23
- [32] Stephanie Strassel. Linguistic Resources for Arabic Handwriting Recognition. In Khalid Choukri and Bente Maegaard, editors, In *Proceedings of the Second International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, April 2009. The MEDAR Consortium. 23
- [33] Dimitra Vergyri and Katrin Kirchhoff. Automatic Diacritization of Arabic for Acoustic Modeling in Speech Recognition. In Ali Farghaly and Karine Megerdoomian, editors, *COLING 2004 Workshop on Computational Approaches to Arabic Script-based Languages*, pages 66–73, Geneva, Switzerland, 2004. DOI: 10.3115/1621804, 24, 37

- [34] S. Ananthakrishnan, S. Narayanan, and S. Bangalore. Automatic Diacritization of Arabic Transcripts for ASR. In *Proceedings of ICON*, Kanpur, India, 2005. 24
- [35] Nizar Habash and Owen Rambow. Arabic Diacritization through Full Morphological Tagging. In *Proceedings of the 8th Meeting of the North American Chapter of the Association for Computational Linguistics/Human Language Technologies Conference (HLT-NAACL07)*, 2007. DOI: 10.3115/1614108.1614122, 24, 86, 88, 104
- [36] Daniel Jurafsky and James H. Martin. *Speech and Language Processing*. Prentice Hall, New Jersey, USA, 2000. 27
- [37] Eugene E.Loos, Susan Anderson, Jr. Dwight H.,Day, Paul C. Jordan, and J. Douglas Wingate. *Glossary of Linguistic Terms*, 2004. 27, 41
- [38] Clive Holes. Modern Arabic: Structures, Functions, and Varieties. *Georgetown Classics in Arabic Language and Linguistics*. Georgetown University Press, 2004. 29, 30, 33, 42, 51, 52, 59, 63
- [39] Janet C. E.Watson. *The Phonology and Morphology of Arabic*. Oxford University Press, 2002. 29, 30, 33
- [40] Nizar Habash. On Arabic and its Dialects. *Multilingual Magazine*, 17(81), 2006. 30
- [41] Fadi Biadsy, Nizar Habash, and Julia Hirschberg. Improving the Arabic Pronunciation Dictionary for Phone and Word Recognition with Linguistically-Based Pronunciation Rules. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 397–405, Boulder, Colorado, June 2009. Association for Computational Linguistics. DOI:10.3115/1620754.1620812, 30, 37
- [42] Y. A. El-Imam. Phonetization of Arabic: Rules and Algorithms. In *Computer Speech and Language* 18, pages 339–373, 2004. DOI: 10.1016/S0885-2308(03)00035-4, 31, 37
- [43] Hany Hassan and Jeffrey Sorensen. An Integrated Approach for Arabic-English Named Entity Translation. In *Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages*, pages 87–93, Ann Arbor, Michigan, June 2005. Association for Computational Linguistics. DOI: 10.3115/1621787.1621803, 36
- [44] Bing Zhao, Nguyen Bach, Ian Lane, and Stephan Vogel. A Log-Linear BlockTransliteration Model based on Bi-Stream HMMs. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Proceedings of the Main Conference*, pages 364–371, Rochester, New York, April 2007. Association for Computational Linguistics. 36

- [45] A. Freeman, S. Condon, and C. Ackerman. Cross Linguistic Name Matching in English and Arabic. In *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*, pages 471–478, New York City, USA, June 2006. Association for Computational Linguistics. DOI: 10.3115/1220835.1220895, 36, 120
- [46] Bassam Haddad and Mustafa Yaseen. Detection and Correction of Non-Words in Arabic: A Hybrid Approach. *International Journal of Computer Processing Of Languages (IJCPOL)*, 2007. DOI: 10.1142/S0219427907001706, 36
- [47] Mohamed Maamouri, Ann Bies, and Seth Kulick. Enhancing the Arabic Treebank: a Collaborative Effort toward New Annotation Guidelines. In European Language Resources Association (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, May 2008. 36, 105
- [48] Chiraz BenOthmane Zribi and Mohammed Ben Ahmed. Efficient Automatic Correction of Misspelled Arabic Words Based on Contextual Information. In *Proceedings of the Knowledge- Based Intelligent Information and Engineering Systems Conference*, Oxford, UK, 2003. 36
- [49] Khaled Shaalan, Amin Allam, and Abdallah Gomah. Towards Automatic Spell Checking for Arabic. In *Conference on Language Engineering*, ELSE, Cairo, Egypt, 2003. 36
- [50] Nizar Habash. Four Techniques for Online Handling of Out-of-Vocabulary Words in Arabic-English Statistical Machine Translation. In *Proceedings of ACL-08: HLT, Short Papers*, pages 57–60, Columbus, Ohio, June 2008. Association for Computational Linguistics. DOI: 10.3115/1557690.1557706, 36, 120, 121
- [51] Nizar Habash and Owen Rambow. Arabic Tokenization, Part-of-Speech Tagging and Morphological Disambiguation in One Fell Swoop. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 573–580, Ann Arbor, Michigan, June 2005. Association for Computational Linguistics. DOI: 10.3115/1219840.1219911, 36, 66, 80, 86, 88, 92
- [52] Linguistic Data Consortium. Egyptian Colloquial Arabic Lexicon. LDC catalog number LDC99L22, ISBN 1-58563-155-8, 1999. 36, 68
- [53] David Graff, Tim Buckwalter, Hubert Jin, and Mohamed Maamouri. Lexicon Development for Varieties of Spoken Colloquial Arabic. In *LREC 2006: Fifth International Conference on Language Resources and Evaluation*, pages 999–1004, Genova, Italy, 2006. 36
- [54] Mona Diab, Nizar Habash, Owen Rambow, Mohamed Altantawy, and Yassine Benajiba. COLABA: Arabic Dialect Annotation and Processing. In *Proceedings of the seventh International Conference on Language Resources and Evaluation (LREC)*, Valletta, Malta, 2010. 36
- [55] Fred Jelinek. Large Vocabulary Continuous Speech Recognition. Technical report, CLSP, Johns Hopkins University, Baltimore, MD, 1997. Summer Research Workshop Technical Reports. 37

- [56] Katrin Kirchhoff, Jeff Bilmes, Sourin Das, Nicolae Duta, Melissa Egan, Gang Ji, Feng He, John Henderson, Daben Liu, Mohamed Noamany, Pat Schone, Richard Schwartz, and Dimitra Vergyri. Novel Approaches to Arabic Speech Recognition: Report from the 2002 Johns-Hopkins Summer Workshop. In *Proceedings of ICASSP 2003*, 2003. DOI: 10.1109/ICASSP.2003.1198788, 37
- [57] Katrin Kirchhoff, Dimitra Vergyri, Jeff A. Blimes, Kevin Duh, and Andreas Stolcke. Morphology-based Language Modeling for Conversational Arabic Speech Recognition. *Computer Speech and Language*, 20:589–608, 2006. DOI: 10.1016/j.csl.2005.10.001, 37
- [58] D.Vergyri,A. Mandal,W.Wang,A. Stolcke, J. Zheng,M.Graciarena,D. Rybach, C. Gollan, R. Schlüter, K. Kirchhoff, A. Faria, and N. Morgan. Development of the SRI/Nightingale Arabic ASR System. In *Proceedings of Interspeech 2008*, 2008. 37
- [59] R. Sproat, editor. *Multilingual Text-to-Speech Synthesis: The Bell Labs Approach*. Kluwer, Boston, MA, 1997. 37
- [60] M.Afify,R. Sarikaya, H.Kuo, L. Besacier, and Y.Gao. On the Use of Morphological Analysis for Dialectal Arabic Speech Recognition. In *Proceedings of Interspeech 2006*, Pittsburgh PA., 2006. 37
- [61] F. Diehl, M.J.F. Gales, M. Tomalin, and P.C.Woodland. Morphological Analysis and Decomposition for Arabic Speech-to-Text Systems. In *Proceedings of InterSpeech*, 2009. 37, 69
- [62] Roger Hsiao, Ashish Venugopal, Thilo Köhler, Ying Zhang, Paisarn Charoenpornasawat, Andreas Zollmann, Stephan Vogel, Alan WBlack, Tanja Schultz, and Alex Waibel. Optimizing Components for Handheld Two-way Speech Translation for an English-Iraqi Arabic System. In *INTERSPEECH*, Pittsburgh, PA, 2006. 37
- [63] Fawzi Alorfi. Automatic Identification of Arabic Dialects Using Hidden Markov Models. PhD thesis, University of Pittsburgh, 2008. 37
- [64] Fadi Biadsy, Julia Hirschberg, and Nizar Habash. Spoken Arabic Dialect Identification Using Phonotactic Modeling. In *Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages*, pages 53–61, Athens, Greece, March 2009. Association for Computational Linguistics. DOI: 10.3115/1621774.1621784, 37
- [65] Fadi Biadsy and Julia Hirschberg. Using Prosody and Phonotactics in Arabic Dialect Identification. In *Proceedings of Interspeech*, Brighton, UK, 2009. 37
- [66] Fadi Biadsy, Andrew Rosenberg, Rolf Carlson, Julia Hirschberg, and Eva Strangert. A Cross-Cultural Comparison of American, Palestinian, and Swedish Perception of Charismatic Speech. In *Speech Prosody*, Campinas, Brazil, 2008. 37
- [67] Otakar Smrž. Functional Arabic Morphology. Formal System and Implementation. PhD thesis, Charles University in Prague, Prague, Czech Republic, 2007. 39, 67, 69, 72, 75, 84, 107.

- [68] Georges Bohas. *Matrices, Étymons, Racines: Éléments d'une théorie lexicographique du vocabulaire arabe*. Peeters, Leuven, 1997. 41
- [69] Z. Harris. Linguistic structure of Hebrew. *Journal of the American Oriental Society*, 62:143–67, 1941. DOI: 10.2307/594501, 43
- [70] John J. McCarthy. A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry*, 12:373–418, 1981. 43
- [71] Mohamed Maamouri and Ann Bies. Developing an Arabic Treebank: Methods, Guidelines, Procedures, and Tools. In *Proceedings of the COLING 2004 Workshop on Computational Approaches to Arabic Script-based Languages*, pages 2–9, 2004. DOI: 10.3115/1621804.1621808, 47, 105
- [72] Mark W. Cowell. *A Reference Grammar of Syrian Arabic*. Georgetown University Press, 1964. 50
- [73] Wallace Erwin. *A Short Reference Grammar of Iraqi Arabic*. Georgetown University Press, 1963. 50
- [74] Ernest T. Abdel-Massih, Zaki N. Abdel-Malek, and El-Said M. Badawi. *A Reference Grammar of Egyptian Arabic*. Georgetown University Press, 1979. 50
- [75] Richard Harrell. *A Short Reference Grammar of Moroccan Arabic*. Georgetown University Press, 1962. 50
- [76] Eckehard Schulz. *A Student Grammar of Modern Standard Arabic*. Cambridge University Press, New York, 2005. 51, 63, 93
- [77] Ron Buckley. *Modern Literary Arabic: A Reference Grammar*. Librairie du Liban, 2004. 52, 57, 61, 63, 93
- [78] William Wright. *A Grammar of the Arabic Language*. Cambridge University Press, reprint of third revised edition, 1991. Translated from the German of Caspari and edited with numerous additions and corrections by W. Wright, revised by W. Robertson Smith and M. J. de Goeje, preface and addenda et corrigenda by Pierre Cachia. 55, 63, 93
- [79] Nizar Habash, Ryan Gabbard, Owen Rambow, Seth Kulick, and Mitch Marcus. Determining Case in Arabic: Learning Complex Linguistic Behavior Requires Complex Linguistic Features. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 1084–1092, 2007. 57, 104, 111
- [80] Nizar Habash and Owen Rambow. MAGEAD: A Morphological Analyzer and Generator for the Arabic Dialects. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 681–688, Sydney, Australia, July 2006. Association for Computational Linguistics. DOI: 10.3115/1220175.1220261, 59, 67, 69, 75.
- [81] Karin C Ryding. *A Reference Grammar of Modern Standard Arabic. Reference Grammars*. Cambridge University Press, New York, 2006. 63, 93.

- [82] Mohamed Maamouri, Ann Bies, Sondos Krouna, Fatma Gaddeche, and Basma Bouziri. Penn Arabic Treebank Guidelines. *Linguistic Data Consortium*, 2009. 63, 80, 93, 104, 105, 106
- [83] Nizar Habash and Fatiha Sadat. Arabic Preprocessing Schemes for Statistical Machine Translation. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 49–52, New York City, USA, June 2006. Association for Computational Linguistics. DOI: 10.3115/1614049.1614062, 66, 76, 77, 88, 89, 121
- [84] Muhammed Aljlal and Ophir Frieder. On Arabic Search: Improving the Retrieval Effectiveness via a Light Stemming Approach. In *Proceedings of ACM Eleventh Conference on Information and Knowledge Management*, Mclean, VA, pages 340–347, 2002. DOI: 10.1145/584792.584848, 67
- [85] Imad Al-Sughayer and Ibrahim Al-Kharashi. Arabic Morphological Analysis Techniques: A Comprehensive Survey. *Journal of the American Society for Information Science and Technology*, 55(3):189–213, 2004. DOI: 10.1002/asi.10368, 67
- [86] Kenneth Beesley. Arabic Finite-State Morphological Analysis and Generation. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING-96)*, pages 89–94, Copenhagen, Denmark, 1996. DOI: 10.3115/992628.992647, 67, 69, 73
- [87] Mohamed Attia Mohamed Elaraby Ahmed. A Large-Scale Computational Processor of the Arabic Morphology and Applications. Master's thesis, Faculty of Engineering, Cairo University, 2000. 67
- [88] Tim Buckwalter. Buckwalter Arabic Morphological Analyzer Version 1.0, 2002. *Linguistic Data Consortium*, University of Pennsylvania. LDC Catalog No.:LDC2002L49. 67, 70, 71, 75, 110
- [89] Kareem Darwish. Building a Shallow Morphological Analyzer in One Day. In *Proceedings of the workshop on Computational Approaches to Semitic Languages in the 40th Annual Meeting of the Association for Computational Linguistics (ACL-02)*, pages 47–54, Philadelphia, PA, USA, 2002. DOI: 10.3115/1118637.1118643, 67, 68
- [90] Jim Yaghi and Sane Yagi. Systematic Verb Stem Generation for Arabic. In *COLING 2004 Computational Approaches to Arabic Script-based Languages*, pages 23–30, Geneva, Switzerland, 2004. DOI: 10.3115/1621804.1621812, 67
- [91] George Kiraz. Multi-tape Two-level Morphology: A Case study in Semitic Non-Linear Morphology. In *Proceedings of Fifteenth International Conference on Computational Linguistics (COLING-94)*, pages 180–186, Kyoto, Japan, 1994. DOI: 10.3115/991886.991917, 67
- [92] Nizar Habash, Owen Rambow, and George Kiraz. Morphological Analysis and Generation for Arabic Dialects. In *Proceedings of the Workshop on Computational Approaches to Semitic Languages at 43rd Meeting of the Association for Computational Linguistics (ACL'05)*, pages 17–24, Ann Arbor, Michigan, 2005. DOI: 10.3115/1621787.1621791, 67, 75

- [93] Mohammed Attia. An Ambiguity-Controlled Morphological Analyzer for Modern Standard Arabic Modelling Finite State Networks. In *The Challenge of Arabic for NLP/MT Conference*, London, 2006. The British Computer Society. 67
- [94] Violetta Cavalli-Sforza, Abdelhadi Souidi, and Teruko Mitamura. Arabic Morphology Generation Using a Concatenative Strategy. In *Proceedings of the 6th Applied Natural Language Processing Conference (ANLP 2000)*, pages 86–93, Seattle, Washington, USA, 2000. 67
- [95] Abdelhadi Souidi, Violetta Cavalli-Sforza, and Abderrahim Jamari. A Computational Lexeme-Based Treatment of Arabic Morphology. In *Proceedings of the Arabic Natural Language Processing Workshop, Conference of the Association for Computational Linguistics (ACL 2001)*, pages 50–57, Toulouse, France, 2001. 67, 69
- [96] Nizar Habash. Large Scale Lexeme Based Arabic Morphological Generation. In *Proceedings of Traitement Automatique des Langues Naturelles (TALN-04)*, pages 271–276, 2004. Fez, Morocco. 67, 69, 71
- [97] Nizar Habash. Arabic Morphological Representations for Machine Translation. In A. van den Bosch and A. Souidi, editors, *Arabic Computational Morphology: Knowledge-based and Empirical Methods*. Springer, 2007. 67, 72, 88, 89
- [98] Franz Josef Och. Google System Description for the 2005 NIST MT Evaluation. In *MT Eval Workshop (unpublished talk)*, 2005. 68, 78, 121
- [99] Ibrahim Badr, Rabih Zbib, and James Glass. Segmentation for English-to-Arabic Statistical Machine Translation. In *Proceedings of ACL-08: HLT, Short Papers*, pages 153–156, Columbus, Ohio, June 2008. Association for Computational Linguistics. DOI: 10.3115/1557690.1557732, 69, 77, 79, 89, 122, 124
- [100] George Anton Kiraz. Multi-Tiered Nonlinear Morphology Using Multi-Tape Finite Automata: A Case Study on Syriac and Arabic. *Computational Linguistics*, 26(1):77–105, 2000. DOI: 10.1162/089120100561647, 73
- [101] Mohamed Altantawy, Nizar Habash, Owen Rambow, and Ibrahim Saleh. Morphological Analysis and Generation of Arabic Nouns: A Morphemic Functional Approach. In *Proceedings of the seventh International Conference on Language Resources and Evaluation (LREC)*, Valletta, Malta, 2010. 75
- [102] Viktor Bielický and Otakar Smrž. Enhancing the ElixirFM Lexicon with Verbal Valency Frames. In Khalid Choukri and Bente Maegaard, editors, In *Proceedings of the Second International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, April 2009. The MEDAR Consortium. 75
- [103] Markus Forsberg and Aarne Ranta. Functional Morphology. In *Proceedings of the Ninth ACM SIGPLAN International Conference on Functional Programming, ICFP 2004*, pages 213–223. ACM Press, 2004. DOI: 10.1145/1016850.1016879, 75.

- [104] Leah S. Larkey, Lisa Ballesteros, and Margaret E. Connell. Arabic Computational Morphology: Knowledge-based and Empirical Methods, chapter *Light Stemming for Arabic Information Retrieval*. Springer Netherlands, kluwer/springer edition, 2007. 77, 90
- [105] Fatiha Sadat and Nizar Habash. Combination of Arabic Preprocessing Schemes for Statistical Machine Translation. In *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*, pages 1–8, Sydney, Australia, July 2006. Association for Computational Linguistics. DOI: 10.3115/1220175.1220176, 77, 79, 88, 89, 121
- [106] Hassan Al-Haj and Alon Lavie. The Impact of Arabic Morphological Segmentation on Broad-coverage English-to-Arabic Statistical Machine Translation. In *Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA)*, Denver, Colorado, 2010. 77, 122, 124
- [107] Jakob Elming and Nizar Habash. Combination of Statistical Word Alignments Based on Multiple Preprocessing Schemes. In *Human Language Technologies 2007: The Conference of the North American Chapter of the Association for Computational Linguistics; Companion Volume, Short Papers*, pages 25–28, Rochester, New York, April 2007. Association for Computational Linguistics. DOI: 10.3115/1614108.1614115, 77, 89, 121
- [108] Young-Suk Lee. Morphological Analysis for Statistical Machine Translation. In *Proceedings of the 5th Meeting of the North American Chapter of the Association for Computational Linguistics/ Human Language Technologies Conference (HLT-NAACL04)*, pages 57–60, Boston, MA, 2004. DOI: 10.3115/1613984.1613999, 79, 121
- [109] Mona Diab, Kadri Hacioglu, and Daniel Jurafsky. Automatic Tagging of Arabic Text: From Raw Text to Base Phrase Chunks. In *Proceedings of the 5th Meeting of the North American Chapter of the Association for Computational Linguistics/Human Language Technologies Conference (HLT-NAACL04)*, pages 149–152, Boston, MA, 2004. DOI: 10.3115/1613984.1614022, 79, 80, 89
- [110] Yuval Marton, Nizar Habash, and Owen Rambow. Improving Arabic Dependency Parsing with Lexical and Inflectional Morphological Features. In *Proceedings of the NAACL HLT 2010 First Workshop on Statistical Parsing of Morphologically-Rich Languages*, pages 13–21, Los Angeles, CA, USA, June 2010. Association for Computational Linguistics. 79, 83, 111, 112
- [111] Mona Diab. Towards an Optimal POS tag set for Modern Standard Arabic Processing. In *Proceedings of Recent Advances in Natural Language Processing (RANLP)*, Borovets, Bulgaria, 2007. 79, 83, 90
- [112] Nizar Habash and Ryan Roth. CATiB: The Columbia Arabic Treebank. In *Proceedings of the ACL-IJCNLP 2009 Conference, Short Papers*, pages 221–224, Suntec, Singapore, August 2009. Association for Computational Linguistics. DOI: 10.3115/1667583.1667651, 79, 83, 104, 108, 112

- [113] Shereen Khoja. APT: Arabic Part-of-Speech Tagger. In *Proceedings of Student Research Workshop at NAACL 2001*, pages 20–26, Pittsburgh, 2001. Association for Computational Linguistics. 80, 84
- [114] Jan Hajič, Otakar Smrž, Tim Buckwalter, and Hubert Jin. Feature-based Tagger of Approximations of Functional Arabic Morphology. In Ma. Antonia Martí Montserrat Civit, Sandra Kübler, editor, *Proceedings of Treebanks and Linguistic Theories (TLT)*, pages 53–64, Barcelona, Spain, 2005. 80, 84, 107
- [115] Noah Smith, David Smith, and Roy Tromble. Context-Based Morphological Disambiguation with Random Fields. In *Proceedings of the 2005 Conference on Empirical Methods in Natural Language Processing (EMNLP05)*, pages 475–482, Vancouver, Canada, 2005. DOI: 10.3115/1220575.1220635, 80
- [116] Ryan Roth, Owen Rambow, Nizar Habash, Mona Diab, and Cynthia Rudin. Arabic Morphological Tagging, Diacritization, and Lemmatization Using Lexeme Models and Feature Ranking. In *Proceedings of ACL-08: HLT, Short Papers*, pages 117–120, Columbus, Ohio, June 2008. Association for Computational Linguistics. DOI: 10.3115/1557690.1557721, 80, 86, 88, 104
- [117] Mona Diab, Kadri Hacioglu, and Daniel Jurafsky. Arabic Computational Morphology: Knowledge-based and Empirical Methods, chapter *Automated Methods for Processing Arabic Text: From Tokenization to Base Phrase Chunking*. Springer Netherlands, kluwer/springer edition, 2007. 80, 89, 90, 112
- [118] Mitchell M. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a Large Annotated Corpus of English: The Penn Treebank. *Computational Linguistics*, 19.2:313–330, June 1993. 80, 105
- [119] Seth Kulick, Ryan Gabbard, and Mitch Marcus. Parsing the Arabic Treebank: Analysis and Improvements. In *Proceedings of the Treebanks and Linguistic Theories Conference*, pages 31–42, Prague, Czech Republic, 2006. 80, 82, 104, 111, 112
- [120] Mona Diab. Improved Arabic Base Phrase Chunking with a New Enriched POS Tag Set. In *Proceedings of the 2007 Workshop on Computational Approaches to Semitic Languages: Common Issues and Resources*, pages 89–96, Prague, Czech Republic, June 2007. Association for Computational Linguistics. DOI: 10.3115/1654576.1654592, 82, 90, 112
- [121] Nizar Habash, Reem Faraj, and Ryan Roth. Syntactic Annotation in the Columbia Arabic Treebank. In *Proceedings of MEDAR International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, 2009. 83, 104, 108
- [122] Shereen Khoja, Roger Garside, and Gerry Knowles. A tagset for the morphosyntactic tagging of Arabic. In *Proceedings of Corpus Linguistics 2001*, pages 341–353, Lancaster, UK, 2001. 84
- [123] Otakar Smrž and Petr Zemánek. Sherds from an Arabic Treebanking Mosaic. *Prague Bulletin of Mathematical Linguistics*, 78:63–76, 2002. 84.

- [124] Marta R. Costa-jussà, Josep M. Crego, Adrià de Gispert, Patrik Lambert, Maxim Khalilov, José A.R. Fonollosa, José B. Mariño, and Rafael Banchs. TALP Phrase-Based System and TALP System Combination for IWSLT 2006. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 123–129, Kyoto, Japan, 2006. 89
- [125] Josep M. Crego, Adrià de Gispert, Patrik Lambert, Maxim Khalilov, Marta R. Costa-jussà, José B. Mariño, Rafael Banchs, and José A.R. Fonollosa. The TALP Ngram-based SMT System for IWSLT2006. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 116–122, Kyoto, Japan, 2006. 89
- [126] David Vilar, Daniel Stein, Yuqi Zhang, Evgeny Matusov, Arne Mauser, Oliver Bender, Saab Mansour, and Hermann Ney. The RWTH Machine Translation System for IWSLT 2008. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 108–115, Hawaii, USA, 2008. 89
- [127] Mona Diab, Mahmoud Ghoneim, and Nizar Habash. Arabic Diacritization in the Context of Statistical Machine Translation. In *Proceedings of Machine Translation Summit (MT-Summit)*, Copenhagen, Denmark, 2007. 89
- [128] Benjamin Farber, Dayne Freitag, Nizar Habash, and Owen Rambow. Improving NER in Arabic Using a Morphological Tagger. In *Proceedings of the Language Resources and Evaluation Conference (LREC)*, Marrakech, Morocco, 2008. 117
- [129] Yassine Benajiba, Mona Diab, and Paolo Rosso. Arabic Named Entity Recognition using Optimized Feature Sets. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing*, pages 284–293, Honolulu, Hawaii, October 2008. Association for Computational Linguistics. DOI: 10.3115/1613715.1613755, 89, 90, 112, 117
- [130] Josep M. Crego and Nizar Habash. Using Shallow Syntax Information to Improve Word Alignment and Reordering for SMT. In *Proceedings of the Third Workshop on Statistical Machine Translation*, pages 53–61, Columbus, Ohio, June 2008. Association for Computational Linguistics. DOI: 10.3115/1626394.1626401, 90, 112, 123
- [131] Behrang Mohit and Rebecca Hwa. Localization of Difficult-to-Translate Phrases. In *Proceedings of the Second Workshop on Statistical Machine Translation*, pages 248–255, Prague, Czech Republic, June 2007. Association for Computational Linguistics. DOI: 10.3115/1626355.1626392, 90
- [132] Nicolas Stroppa and Andy Way. MATREX: DCU Machine Translation System for IWSLT 2006. In *Proceedings of the International Workshop on Spoken Language Translation*, pages 31–36, Kyoto, Japan, 2006. 90
- [133] David Farwell, Jesús Giménez, Edgar González, Reda Halkoum, Horacio Rodríguez, and Mihai Surdeanu. The UPC System for Arabic-to-English Entity Translation. In *Proceedings of ACE 2007*, 2007. 90
- [134] Robert D. Van Valin. *An Introduction to Syntax*. Cambridge University Press, 2001. 93

- [135] Beatrice Santorini and Anthony Kroch. *The Syntax of Natural Language: An Online Introduction using the Trees Program*, 2007. 93
- [136] Mohamed Maamouri, Ann Bies, and Seth Kulick. Creating a Methodology for Large-Scale Correction of Treebank Annotation: The Case of the Arabic Treebank. In *Proceedings of MEDAR International Conference on Arabic Language Resources and Tools*, Cairo, Egypt, 2009. 93, 104, 105
- [137] Otakar Smrž and Jan Hajič. The Other Arabic Treebank: Prague Dependencies and Functions. In Ali Farghaly, editor, *Arabic Computational Linguistics: Current Implementations*. CSLI Publications, 2006. 104
- [138] Otakar Smrž, Viktor Bieličký, Iveta Kouřilová, Jakub Kráčmar, Jan Hajič, and Petr Zemánek. Prague Arabic Dependency Treebank: A Word on the Million Words. In *Proceedings of the Workshop on Arabic and Local Languages (LREC 2008)*, pages 16–23, Marrakech, Morocco, 2008. 104, 107, 117
- [139] Mohamed Maamouri and Christopher Cieri. Resources for Natural Language Processing at the Linguistic Data Consortium. In *Proceedings of the International Symposium on Processing of Arabic*, pages 125–146, Manouba, Tunisia, 2002. 105
- [140] Mohamed Maamouri, Ann Bies, Tim Buckwalter, Mona Diab, Nizar Habash, Owen Rambow, and Dalila Tabessi. Developing and Using a Pilot Dialectal Arabic Treebank. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation, LREC'06*, Genoa, Italy, 2006. 105, 111
- [141] Petr Sgall, Eva Hajičová, and Jarmila Panevová. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. D. Reidel & Academia, 1986. 106
- [142] Jan Hajič, Barbora Hladká, and Petr Pajas. The Prague Dependency Treebank: Annotation Structure and Support. In *Proceedings of the IRCS Workshop on Linguistic Databases*, pages 105–114, Philadelphia, 2001. University of Pennsylvania. 106
- [143] Otakar Smrž. ElixirFM — Implementation of Functional Arabic Morphology. In *Proceedings of the 2007 Workshop on Computational Approaches to Semitic Languages: Common Issues and Resources*, pages 1–8, Prague, Czech Republic, June 2007. ACL. DOI: 10.3115/1654576.1654578, 106, 107
- [144] Zdeněk Žabokrtský and Otakar Smrž. Arabic Syntactic Trees: from Constituency to Dependency. In *Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics (EACL'03) – Research Notes*, Budapest, Hungary, 2003. DOI: 10.3115/1067737.1067779, 107
- [145] Jan Hajič, Otakar Smrž, Petr Zemánek, Jan Šnaidauf, and Emanuel Beška. Prague Arabic Dependency Treebank: Development in Data and Tools. In *NEMLAR International Conference on Arabic Language Resources and Tools*, pages 110–117. ELDA, 2004. 107 .

- [146] Joakim Nivre, Johan Hall, Jens Nilsson, Atanas Chanev, Gulsen Eryigit, Sandra Kubler, Svetoslav Marinov, and Erwin Marsi. MaltParser: A Language-independent System for Data-driven Dependency Parsing. *Natural Language Engineering*, 13(2):95–135, 2007. DOI: 10.1017/S1351324906004505, 107, 112
- [147] Fei Xia, Owen Rambow, Rajesh Bhatt, Martha Palmer, and Dipti Misra Sharma. Towards a Multi-Representational Treebank. In *Proceedings of Treebanks and Linguistic Theories (TLT 7)*, Groningen, Netherlands, 2009. 109
- [148] David Chiang, Mona Diab, Nizar Habash, Owen Rambow, and Safiullah Shareef. Parsing Arabic Dialects. In *Proceedings of the European Chapter of ACL (EACL)*, 2006. 111, 112
- [149] Kais Dukes and Tim Buckwalter. A Dependency Treebank of the Quran using Traditional Arabic Grammar. In *Proceedings of the 7th international conference on Informatics and Systems (INFOS 2010)*, Cairo, Egypt, 2010. 112
- [150] Lamia Tounsi, Mohammed Attia, and Josef van Genabith. Automatic Treebank-Based Acquisition of Arabic LFG Dependency Structures. In *Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages*, pages 45–52, Athens, Greece, 2009. DOI: 10.3115/1621774.1621783, 112
- [151] Martha Palmer, Olga Babko-Malaya, Ann Bies, Mona Diab, Mohamed Maamouri, Aous Mansouri, and Wajdi Zaghouani. A Pilot Arabic Propbank. In *Proceedings of LREC*, Marrakech, Morocco, May 2008. 112, 114
- [152] Eduard Hovy, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. OntoNotes: The 90% Solution. In *NAACL '06: Proceedings of the Human Language Technology Conference of the NAACL*, Companion Volume: Short Papers on XX, pages 57–60, Morristown, NJ, USA, 2006. Association for Computational Linguistics. DOI: 10.3115/1614049.1614064, 112, 117
- [153] Daniel Bikel. Design of a Multi-lingual, Parallel-processing Statistical Parsing Engine. In *Proceedings of International Conference on Human Language Technology Research (HLT)*, pages 24–27, 2002. DOI: 10.3115/1289189.1289191, 112
- [154] Ryan Gabbard and Seth Kulick. Construct State Modification in the Arabic Treebank. In *Proceedings of ACL-08: HLT*, Short Papers, pages 209–212, Columbus, Ohio, June 2008. Association for Computational Linguistics. DOI: 10.3115/1557690.1557750, 112
- [155] Dan Klein and Christopher D. Manning. Accurate Unlexicalized Parsing. In *Proceedings of the 41st Meeting of the Association for Computational Linguistics (ACL'03)*, 2003. DOI: 10.3115/1075096.1075150, 112.

- [156] Spence Green, Conal Sathi, and Christopher D. Manning. NP Subject Detection in Verbinitial Arabic Clauses. In *Proceedings of the Third Workshop on Computational Approaches to Arabic Script-based Languages (CAASL3)*, 2009. 112, 123
- [157] Michael C. McCord and Violetta Cavalli-Sforza. An Arabic Slot Grammar Parser. In *Proceedings of the 2007 Workshop on Computational Approaches to Semitic Languages*, pages 81–88, Morristown, NJ, USA, 2007. Association for Computational Linguistics. DOI: 10.3115/1654576.1654591, 112
- [158] Eman Othman, Khaled Shaalan, and Ahmed Rafea. A Chart Parser for Analyzing Modern Standard Arabic Sentence. In *Proceedings of the MT Summit IX Workshop on Machine Translation for Semitic Languages: Issues and Approaches*, pages 37–44, 2003. 112
- [159] D. Cruse. *Lexical Semantics*. Cambridge University Press, 1986. 113
- [160] R. Jackendoff. *Semantic Structures*. MIT Press, Boston, Mass, 1990. 113
- [161] G. Chierchia and S. McConell-Ginet. *Meaning and Grammar: An Introduction to Semantics*. The MIT Press, Cambridge, MA, 1990. 113
- [162] M. Palmer, D. Gildea, and P. Kingsbury. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106, 2005. DOI: 10.1162/0891201053630264, 114
- [163] Collin F. Baker, Charles J. Fillmore, and John B. Lowe. The Berkeley FrameNet Project. In *COLING-ACL '98: Proceedings of the Conference*, held at the University of Montréal, pages 86–90, 1998. DOI: 10.3115/980845.980860, 114
- [164] Nianwen Xue and Martha Palmer. Adding Semantic Roles to the Chinese Treebank. *Nat. Lang. Eng.*, 15(1):143–172, 2009. DOI: 10.1017/S1351324908004865, 114
- [165] Mona Diab, Alessandro Moschitti, and Daniele Pighin. Semantic Role Labeling Systems for Arabic using Kernel Methods. In *Proceedings of ACL-08: HLT*, pages 798–806, Columbus, Ohio, June 2008. Association for Computational Linguistics. 115
- [166] Mona Diab, Musa Alkhalifa, Sabry ElKateb, Christiane Fellbaum, Aous Mansouri, and Martha Palmer. SemEval-2007 Task 18: Arabic Semantic Labeling. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 93–98, Prague, Czech Republic, June 2007. Association for Computational Linguistics. DOI: 10.3115/1621474.1621491, 115, 116
- [167] Christiane Fellbaum. *WordNet: An Electronic Lexical Database*. MIT Press, 1998. <http://www.cogsci.princeton.edu/~wn> [2000, September 7]. 116
- [168] Piek Vossen. *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Kluwer Academic Publishers, Dordrecht, 1998. 116
- [169] W. Black, S. ElKateb, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, and C. Fellbaum. Introducing the Arabic WordNet Project. In *Proceedings of the Third International WordNet Conference*, Jeju Island, Korea, 2006. 116

- [170] S. Elkateb, W. Black, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, and C. Fellbaum. Building a WordNet for Arabic. In *Proceedings of The fifth international conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, 2006. 116
- [171] G. Doddington, A. Mitchell, M. Przybocki, L. Ramshaw, S. Strassel, and R. Weischedel. The Automatic Content Extraction (ACE) Program—Tasks, Data, and Evaluation. In *Proceedings of LREC 2004*, pages 837–840, 2004. 117
- [172] Imed Zitouni, Jeffrey Sorensen, Xiaoqiang Luo, and Radu Florian. The impact of morphological stemming on Arabic mention detection and coreference resolution. In *Proceedings of the ACL Workshop on Computational Approaches to Semitic Languages*, pages 63–70, Ann Arbor, Michigan, June 2005. 117
- [173] Yassine Benajiba and Imed Zitouni. Morphology based Segmentation Combination for Arabic Mention Detection. *Special issue on Arabic Natural Processing of the ACM Transactions on Asian Language Information Processing (TALIP)*, 2010. DOI: 10.1145/1644879.1644883, 117
- [174] Mona T. Diab. An Unsupervised Approach for Bootstrapping Arabic Sense Tagging. In Ali Farghaly and Karine Megerdooian, editors, *COLING 2004 Computational Approaches to Arabic Script-based Languages*, pages 43–50, Geneva, Switzerland, August 28th 2004. COLING. 117
- [175] David Farwell, Stephen Helmreich, Florence Reeder, Bonnie Dorr, Nizar Habash, Eduard Hovy, Lori Levin, Keith Miller, Teruko Mitamura, Owen Rambow, and Advait Siddharthan. Interlingual Annotation of Multilingual Text Corpus. In *Proceedings of the NAACL/HLT Workshop: New Frontiers in Corpus Annotation*, 2004. 117
- [176] Nizar Habash, Clinton Mah, Sabiha Imran, Randy Calistri-Yeh, and Páiraic Sheridan. Design, Construction and Validation of an Arabic-English Conceptual Interlingua for Cross-lingual Information Retrieval. In *LREC-2006: Fifth International Conference on Language Resources and Evaluation*, pages 107–112, Genoa, Italy, 2006. 118
- [177] Bonnie J. Dorr, Pamela W. Jordan, and John W. Benoit. A Survey of Current Research in Machine Translation. In M. Zelkowitz, editor, *Advances in Computers*, 49, pages 1–68. Academic Press, London, 1999. 119
- [178] Franz Josef Och and Hermann Ney. A Systematic Comparison of Various Statistical Alignment Models. *Computational Linguistics*, 29(1):19–52, 2003. DOI: 10.1162/089120103321337421, 119
- [179] Philipp Koehn. Pharaoh: a Beam Search Decoder for Phrase-based Statistical Machine Translation Models. In *Proceedings of the Association for Machine Translation in the Americas*, pages 115–124, 2004. 119
- [180] Jan Hajič, Jan Hric, and Vladislav Kubon. Machine Translation of Very Close Languages. In *Proceedings of the 6th Applied Natural Language Processing Conference (ANLP'2000)*, pages 7–12, Seattle, 2000. DOI: 10.3115/974147.974149, 120

- [181] Andreas Zollmann, Ashish Venugopal, and Stephan Vogel. Bridging the Inflection Morphology Gap for Arabic Statistical Machine Translation. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 201–204, New York City, USA, 2006. Association for Computational Linguistics. DOI: 10.3115/1614049.1614100, 121
- [182] Jason Riesa and David Yarowsky. Minimally Supervised Morphological Segmentation with Applications to Machine Translation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA06)*, pages 185–192, Cambridge, MA, 2006. 121
- [183] Anas El Isbihani, Shahram Khadivi, Oliver Bender, and Hermann Ney. Morpho-syntactic Arabic Preprocessing for Arabic to English Statistical Machine Translation. In *Proceedings on the Workshop on Statistical Machine Translation*, pages 15–22, New York City, June 2006. Association for Computational Linguistics. DOI: 10.3115/1654650.1654654, 121
- [184] Nizar Habash. Syntactic Preprocessing for Statistical MT. In *Proceedings of the Machine Translation Summit (MT SUMMIT XI)*, Copenhagen, Denmark, 2007. 123
- [185] Nizar Habash, Bonnie Dorr, and Christof Monz. Challenges in Building an Arabic-English GHMT System with SMT Components. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas (AMTA06)*, pages 56–65, Cambridge, MA, 2006. 123
- [186] Steve DeNeefe and Kevin Knight. Synchronous Tree Adjoining Machine Translation. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 727–736, Singapore, August 2009. Association for Computational Linguistics. DOI: 10.3115/1699571.1699607, 123
- [187] Marine Carpuat, Yuval Marton, and Nizar Habash. Improving Arabic-to-English Statistical Machine Translation by Reordering Post-Verbal Subjects for Alignment. In *Proceedings of the ACL 2010 Conference, Short Papers*, pages 178–183, Uppsala, Sweden, July 2010. Association for Computational Linguistics. 123
- [188] Ibrahim Badr, Rabih Zbib, and James Glass. Syntactic Phrase Reordering for English-to-Arabic Statistical Machine Translation. In *Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009)*, pages 86–93, Athens, Greece, March 2009. Association for Computational Linguistics. DOI: 10.3115/1609067.1609076, 123, 124
- [189] Jakob Elming and Nizar Habash. Syntactic Reordering for English-Arabic Phrase-Based Machine Translation. In *Proceedings of the EACL 2009 Workshop on Computational Approaches to Semitic Languages*, pages 69–77, Athens, Greece, March 2009. Association for Computational Linguistics. DOI: 10.3115/1621774.1621786, 123.

- [190] Saša Hasan, Anas El Isbihani, and Hermann Ney. Creating a Large-Scale Arabic to French Statistical Machine Translation System. In *Proceedings of Language Resources and Evaluation Conference (LREC)*, pages 855–858, Genoa, Italy, 2006. 124
- [191] Nizar Habash and Jun Hu. Improving Arabic-Chinese Statistical Machine Translation using English as Pivot Language. In *Proceedings of the Fourth Workshop on Statistical Machine Translation*, pages 173–181, Athens, Greece, March 2009. Association for Computational Linguistics. DOI: 10.3115/1626431.1626467, 124
- [192] Mossab Al-Hunaity, Bente Maegaard, and Dorte Hansen. Using English as a Pivot Language to Enhance Danish-Arabic Statistical Machine Translation. In *Workshop on Language Resources and Human Language Technology for Semitic Languages in the Language Resources and Evaluation Conference (LREC)*, Valletta, Malta, 2010. 124
- [193] Reshef Shilon, Nizar Habash, Alon Lavie, and ShulyWintner. Machine Translation between Hebrew and Arabic: Needs, Challenges and Preliminary Solutions. In *Proceedings of the Student Research Workshop in the Conference of the Association for Machine Translation in the Americas (AMTA)*, Denver, Colorado, 2010. 124
- [194] Haytham Alsharaf, Sylviane Cardey, Peter Greenfield, and Yihui Shen. Problems and Solutions in Machine Translation Involving Arabic, Chinese and French. In *Proceedings of the International Conference on Information Technology*, pages 293–297, Las Vegas, Nevada, 2004. DOI: 10.1109/ITCC.2004.1286649, 124
- [195] Mohammed Sharaf. Implications of the Agreement Features in (English to Arabic) Machine Translation. Master's thesis, Al-Azhar University, 2002. 124
- [196] Abdelhadi Soudi, Violetta Cavalli-Sforza, and Abderrahim Jamari. A Prototype English-to-Arabic Interlingua-based MT system. In *Proceedings of the Third International Conference on Language Resources and Evaluation: Workshop on Arabic language resources and evaluation*, Las Palmas, Spain, 2002. 124
- [197] Abdelhadi Soudi. Challenges in the Generation of Arabic from Interlingua. In *Proceedings of Traitement Automatique des Langues Naturelles (TALN-04)*, pages 343–350, 2004. Fez, Morocco. 124
- [198] Azza Abdel-Monem, Khaled Shaalan, Ahmed Rafea, and Hoda Baraka. A Proposed Approach for Generating Arabic from Interlingua in a Multilingual Machine Translation System. In *Proceedings of the 4th Conference on Language Engineering*, pages 197–206, 2003. Cairo, Egypt. 124
- [199] F. Gey and D. Oard. The TREC-2001 Cross-Language Information Retrieval Track: Searching Arabic Using English, French or Arabic Queries. In *The 10th Text Retrieval Conference (TREC-10)*, 2001. 124
- [200] F. Gey and D. Oard. The TREC-2002 Arabic/English Cross-Language Information Retrieval Track. In *The 11th Text Retrieval Conference (TREC-11)*, 2002. 124

- [201] Jinxi Xu, Alexander Fraser, and Ralph Weischedel. Empirical Studies in Strategies for Arabic Retrieval. In *SIGIR '02: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 269–274, New York, NY, USA, 2002. ACM. DOI: 10.1145/564376.564424, 124
- [202] Leah S. Larkey, Lisa Ballesteros, and Margaret E. Connell. Improving Stemming for Arabic Information Retrieval: Light Stemming and Co-occurrence Analysis. In *Proceedings of the 25th Annual International Conference on Research and Development in Information Retrieval (SIGIR 2002)*, Tampere, Finland, pages 275–282, 2002. DOI: 10.1145/564376.564425, 124
- [203] K. Darwish and D. Oard. CLIR Experiments at Maryland for TREC 2002: Evidence Combination for Arabic-English Retrieval. In *The 11th Text Retrieval Conference (TREC-11)*, 2002. 124
- [204] Ramzi Abbès, Joseph Dichy, and Mohamed Hassoun. The Architecture of a Standard Arabic Lexical Database. Some Figures, Ratios and Categories from the DIINAR.1 Source Program. In Ali Farghaly and Karine Megerdumian, editors, *COLING 2004 Computational Approaches to Arabic Script-based Languages*, pages 15–22, Geneva, Switzerland, August 28<sup>th</sup> 2004. COLING. 138.

obeikandi.com

## ثبت المصطلحات

أولاً: عربي - إنجليزي

أ

Coreference	الإحالة
Indicative	إخبارية
Orthographic normalization	الاختزال الهجائي
Particle	أداة
Definite article	أداة التعريف
Verbal progressive particle	أداة الفعل المضارع
Subordinating conjunction	أداة ربط ثانوية
Assimilation	إدغام
Subordination conjunction	أدوات ربط الجملة الثانوية
Prefixation	إسباق
Elision	الإسقاط
Pro-drop	إسقاط الضمير
Demonstrative pronouns	اسم الإشارة

Diminutive form	اسم التصغير
Active participle	اسم الفاعل
Passive participle	اسم المفعول
Relative pronouns	الاسم الموصول
Nominal indefiniteness	الاسم النكرة
Relative pronoun	اسم موصول
Deverbal nouns	الأسماء المشتقة من أفعال
Quantifier nouns	الأسماء المحدودة
Relative clauses	الأسماء الموصولة/عبارات الوصل
Emphatics	الأصوات المفخمة
Phonemic	الأصواتية/علم دراسة الوحدات الصوتية
Genitive	إضافة/ مجرور
Imperfective verbs	أفعال ناقصة
Suffixation	الإلحاق
Homonymy	الألفاظ المتجانسة/ الجناس
Allomorph	الألومورف
Tag	أمانة
Orthography	الإملاء
Enclitics	الإنضائي/ الزوائد في نهاية الكلمة

## ب

Propbank	بنك الأبنية الحملية
----------	---------------------

Prague Arabic dependency Treebank	بنك براغ للتحليل النحوية التبعية للغة العربية / شجرة براغ التبعية للغة العربية
Propositional bank	بنك خبري
Residuals	البواقي
<b>ن</b>	
Part of Speech Tagging	تأشير/وصف أجزاء الكلام / العنونة
Stemming	التجذير/التجذيع
Detokenization	تجميع النص / عدم التقطيع
Label	تسمية
Normalization	تسوية / توحيد
Interdigitated	تشابك
Lemmatization	التشذيب
Verb conjugated	تصريف الأفعال
Inflection	التصريف / الإعراب
Inflectional	تصريفية
Homophones	التطابق اللفظي
Morphological adjustments	التعديلات الصرفية
Speech recognition	التعرف على الكلام / تمييز الأصوات
Named entity recognition	التعرف على الأسماء
Definiteness	التعريف
Normalized tokenization	التقطيع المختزل

Tokenization	تقطيع النص
Orthographic transliteration	تهجئة النقل الكتابي/النقحره
Semantic role labeling	توسيم/عنوانه الأءوار الءلالية
Speech synthesis	ءوليف الكلام

## ث

Triptotes	ءلائي الإعراب (الصرف)/ المصرف
Diptotes	ءنائى الإعراب (الصرف)/ الممنوع من الصرف

## ج

Jussive	الجزم
Nominal sentence	ءمل اسمية
Homographs	ءناس
Homograph	ءنيس ءط/مءترك

## ء

Adverb	ءال
Construct state	ءالة الإءضافة/البناء (المبنى)
Prepositional	ءالة الجر
Case	ءالة النءوية
Glide	ءرف العلة
Preposition	ءرف جر
Prepositional proclitic	ءرف جر زائد

Short vowels	حركات / صوائت قصيرة
Short vowel	حركات / صوائت قصيرة
Radical	الحروف الأصلية
Coordinating conjunction	حروف العطف
Vocalism	حروف العلة / الإعلال

## خ

Proposition	خبر
-------------	-----

## ز

Affix	الزائدة / لاصقة
Aspect	الزمن
Perfective aspect	الزمن المكتمل
Clitic	زوائد
Pronominal enclitic	الزوائد الضميرية في بداية / نهاية الكلمة
Proclitics	زوائد في بداية الكلمة
Affixation	الزيادة / الإلصاق

## ش

Dependency Treebank	الشجرات التبعية
Penn Arabic Treebank	شجرة بنسلفانيا العربية / بنك بنسلفانيا للتحليل النحوية
Subjunctive	الشرط (صيغة الشرط ، صيغة المضارع المنصوب)

Perfective form

شكل سالم

**ص**

Diphthong

الصائت المركب (إدغام)

Morpho-phonemically

الصرف - فنولوجي / المورفونولوجية

Derivational morphology

الصرف الاشتقائي

Inflectional morphology

الصرف التصريفي

Functional morphology

الصرف الوظيفي

Morpho-semantic

الصرف دلالي

Morpho-syntactic

الصرف نحوية

Morpheme

صَرْفَم / وَحْدَةٌ صَرْفِيَّةٌ

Morphophonemic

صرفمي / علم الأصوات الصرفي

Adjective

صفة

Tuple

صفوف علائقية

Phoneme

صَوْتَم (وَحْدَةٌ لَفْظِيَّةٌ)

Mood

صيغة الفعل

Interjection

صيغة تعجب

**ض**

Object pronoun

ضمائر المفعول به

Interrogative pronoun

ضمير استفهام

Pronominal

ضميري

## عم

Nominals	العبارات الاسمية / الكلمات الاسمية
Lexeme	العجم أو وحدة معجمية
Token	علامة
Phonetics	علم الأصوات
Stem-vowel	عين الفعل

## نم

3rd person	الغائب
------------	--------

## ف

Imperative verb	فعل أمر
Perfective verb	فعل ماضي
Decliticization	فك الزوائد

## ك

Transcription	الكتابة الصوتية
Invariables	الكلمات المبنية
Indeclinables	الكلمات المعربة
Lemma	الكلمة المعجمية
Gloss	الكلمة الموضحة / حاشية

## ل

Computational linguistics	اللسانيات الحاسوبية
---------------------------	---------------------

Native language

اللغة القومية

م

String

متوالية

Synsets Synonym Set

المجموعات المترادفة

Base phrase chunker

المحلل السطحي للعبارات / مجزئ العبارات

1st person

المخاطب

Second person

مخاطب

Realizations

المدركات

Geminate

مدغم

Nominative

مرفوع

Subject

المسند إليه / الفاعل

Deverbal

مشتقات فعلية

Verbal-noun

المصدر

Nominative

مضاف

Verbal aspect

المظهر اللفظي

Defective

معتل

Vocalic

معتل

Tagger

معنون

Emphatic

مفخم

3rd person masculine singular

مفرد مذكر غائب

Mapping

مقابلة

Comparative adjectives	مقارنة الصفات
Cliticization	الملحقات
Accusative	منصوب
<b>ن</b>	
Lexical functional grammar	نحو وظيفي معجمي
Syntax	نظم الجملة
Participle	النعث
Transliteration	النقل الكتابي
Indefinite	نكرة
<b>و</b>	
Pattern	وزن

## ثانياً: إنجليزي - عربي

## A

Accusative	منصوب
Active participle	اسم الفاعل
Adjective	صفة
Adverb	حال
Affix	الزائدة/ لاصقة
Affixation	الزيادة/ الإلصاق
Allomorph	الألومورف
Aspect	الزمن
Assimilation	إدغام

## B

Base phrase chunker	المحلل السطحي للعبارات/ مجزئ العبارات
---------------------	---------------------------------------

## C

Case	الحالة النحوية
Clitic	زوائد
Cliticization	الملحقات
Comparative adjectives	مقارنة الصفات
Computational linguistics	اللسانيات الحاسوبية
Construct state	حالة الإضافة/ البناء (المبنى)

Coordinating conjunction حروف العطف

Coreference الإحالة

## D

Decliticization فك الزوائد

Defective معتل

Definite article أداة التعريف

Definiteness التعريف

Demonstrative pronouns اسم الإشارة

Dependency Treebank الشجرات التبعية

Derivational morphology الصرف الاشتقائي

Detokenization تجميع النص / عدم التقطيع

Deverbal nouns الأسماء المشتقة من أفعال

Deverbals مشتقات فعلية

Diminutive form اسم التصغير

Diphthong الصائت المركب (إدغام)

Diptotes ثنائي الإعراب (الصرف) / الممنوع

من الصرف

## E

Elision الإسقاط

Emphatic مفخم

Emphatics الأصوات المفخمة

Enclitics الإِنْصَوَائِي / الزوائد في نهاية الكلمة

**F**

Functional morphology الصرف الوظيفي

**G**

Geminate مدغم

Genitive إضافة / مجرور

Glide حرف العلة

Gloss الكلمة الموضحة / حاشية

**H**

Homograph جنيس خط / مشترك

Homographs جناس

Homonymy الألفاظ المتجانسة / الجناس

Homophones التطابق اللفظي

**I**

Imperative verb فعل أمر

Imperfective verbs أفعال ناقصة

Indeclinables الكلمات المعربة

Indefinite نكرة

Indicative إخبارية

Inflection التصريف / الإعراب

Inflectional	تصرفية*
Inflectional morphology	الصرف التصريفي
Interdigitated	تشابك
Interjection	صيغة تعجب
Interrogative pronoun	ضمير استفهام
Invariables	الكلمات المبنية

**J**

Jussive	الجزم
---------	-------

**L**

Label	تسمية
Lemma	الكلمة المعجمية
Lemmatization	التشذيب
Lexeme	العجم أو وحدة معجمية
Lexical functional grammar	نحو وظيفي معجمي

**M**

Mapping	مقابلة
Mood	صيغة الفعل
Morpheme	صَرْفَم / وَحْدَة صَرْفِيَّة
Morphological adjustments	التعديلات الصرفية
Morphophonemic	صرفمي / علم الأصوات الصرفي

Morpho-phonemically	الصرف – فنولوجي / المورفونولوجية
Morpho-semantic	الصرف دلالي
Morpho-syntactic	الصرف نحوية

**N**

Named entity recognition	التعرف على الأسماء
Native language	اللغة القومية
Nominal indefiniteness	الاسم النكرة
Nominal sentence	جمل اسمية
Nominals	العبارات الاسمية / الكلمات الاسمية
Nominative	مرفوع
Nominative	مضاف
Normalization	تسوية/توحيد
Normalized tokenization	التقطيع المختزل

**O**

Object pronoun	ضمائر المفعول به
Orthographic normalization	الاختزال الهجائي
Orthographic transliteration	تهجئة النقل الكتابي/النقحرة
Orthography	الإملاء

**P**

Part of Speech Tagging	تأشير/وصف أجزاء الكلام / العنونة
------------------------	----------------------------------

Participle	النعته
Particle	أداة
Passive participle	اسم المفعول
Pattern	وزن
Penn Arabic Treebank	شجرة بنسلفانيا العربية / بنك بنسلفانيا للتحاليل النحوية
Perfective aspect	الزمن المكتمل
Perfective form	شكل سالم
Perfective verb	فعل ماضي
Phoneme	صَوْتَم (وَحْدَةٌ لَفْظِيَّةٌ)
Phonemic	الأصواتية / علم دراسة الوحدات الصوتية
Phonetics	علم الأصوات
Prague Arabic dependency Treebank	بنك براغ للتحاليل النحوية التابعة للغة العربية / شجرة براغ التابعة للغة العربية
Prefixation	إِسْبَاقٌ
Preposition	حرف جر
Prepositional	حالة الجر
Prepositional proclitic	حرف جر زائد
Proclitics	زوائد في بداية الكلمة
Pro-drop	إسقاط الضمير
Pronominal	ضميري
Pronominal enclitic	الزوائد الضميرية في بداية / نهاية الكلمة

Propbank	بنك الأبنية الحملية
Proposition	خبر
Propositional bank	بنك خبري

**Q**

Quantifier nouns	الأسماء المعدودة
------------------	------------------

**R**

Radical	الحروف الأصلية
Realizations	المدركات
Relative clauses	الأسماء الموصولة/عبارات الوصل
Relative pronoun	اسم موصول
Relative pronouns	الاسم الموصول
Residuals	البواقي

**S**

Second person	مخاطب
Semantic role labeling	توسيم/عنوانة الأدوار الدلالية
Short vowel	حركات/صوائت قصيرة
Short vowels	حركات/صوائت قصيرة
Speech recognition	التعرف على الكلام/تمييز الأصوات
Speech synthesis	توليف الكلام
Stemming	التجذير/التجذيع

Stem-vowel	عين الفعل
String	متوالية
Subject	المسند إليه / الفاعل
Subjunctive	الشرط (صيغة الشرط ، صيغة المضارع المنصوب)
Subordinating conjunction	أداة ربط ثانوية
Subordination conjunction	أدوات ربط الجملة الثانوية
Suffixation	الإلحاق
Synsets Synonym Set	المجموعات المترادفة
Syntax	نظم الجملة

## T

Tag	أمانة
Tagger	معنون
Token	علامة
Tokenization	تقطيع النص
Transcription	الكتابة الصوتية
Transliteration	النقل الكتابي
Triptotes	ثلاثي الإعراب (الصرف) / المصرف
Tuple	صفوف علائقية

## V

Verb conjugated	تصريف الأفعال
-----------------	---------------

Verbal aspect

المظهر اللفظي

Verbal progressive particle

أداة الفعل المضارع

Verbal-noun

المصدر

Vocalic

معتل

Vocalism

حروف العلة / الإعلال

## كشاف الموضوعات

ألوقرافي ١٢، ٣٠، ٣٥، ٣٩، ٤٠  
أميرة ١٤٧، ١٥٢، ١٥٧، ١٥٩،  
١٦٠، ١٦٢، ١٦٣  
أنظمة التعرف ١٣، ٦٦  
أنظمة فك اللبس الدلالي للعربية ٢٠٦  
إزالة التشكيل ٣٩  
إزالة التطويل ٣٩  
استخلاص المعلومات ٢٠٧، ٢٠٨  
استرجاع المعلومات عبر اللغات ٢٢١،  
٢٦٤  
اسم الجمع ١٠٠، ١٠٤  
الاتجاهية ١١٧، ١٢٢  
الإحالة ١٧٢، ١٩٥، ٢٠٧، ٢٠٨،  
٢٤٨، ٢٨٩، ٢٩٩

## أ

أداة المحلل ٦٥  
أداة أميرة ١٥٢، ١٥٩  
أسماء الأعلام ٣٧، ٤٧، ٦٢، ٦٤،  
١٠٢، ١٤٩، ١٥٥، ١٥٧،  
١٦١، ١٩٧، ٢٦٢  
أشكال الأفعال ٩٠، ٩١، ١٠٣  
أزواج الفونيمات الصوتية ٥١  
ألفبائية ٩، ١١، ١٧  
ألوفورية ١٣٣  
ألوفونات ٤٨، ٥٢، ٧٧  
ألوقراف ١٢، ٣٠، ٣٥، ٣٩، ٤٠

الأنتولوجيا المعجمية ٢٠٥	الإدغام ٥٢ ، ٦٠ ، ٦٣ ، ٨٨ ، ١٠٩ ، ١١١ ، ١١٠
<b>ب</b>	الأدوار الدلالية ١٨٢ ، ٢٠٠ ، ٢٠٤ ، ٢٩٢
باما ١٢٣ ، ١٢٤ ، ١٣٠ ، ١٨٠ ، ٢٥٨ ، ٢٠٨ ، ٢٠٧	الأرقام العربية - الهندية ٢٢ ، ٢٣ ، ٣٥
بكوالتر ٧ ، ٣٦ ، ٤٠ ، ٥٦ ، ١٤١ ، ١٤٢ ، ١٤٣ ، ١٤٧ ، ١٩٣	الازدواج اللغوي ٤
بنك أبنية الحمل ٢٠١ ، ٢٠٢ ، ٢٠٤ ، بنك الأبنية الحملية ١٩٦ ، ٢٠٣ ، ٢٤٧ ، ٢٩١ ، ٣٠٤	الاستخلاص التلقائي للمحتوى ٢٠٨
بنك براغ للتحاليل الشجرية النحوية	الاسم المقصور ١٠٢
التبعية للغة العربية ١٨٦ ، ١٨٧ ، ٢٠٣ ، ٢٠٤ ، ٢٠٩	الاسم المنقوص ١٠٢
بنك بنسلفانيا الشجري للتحاليل النحوية ٨٥ ، ٨٦ ، ٨٨ ، ١٨٦ ، ٢٠١	الإسناد ١٧٣ ، ١٧٤ ، ٢٠١ ، ٢٠٢ ، ٢٩٩ ، ٢٩٤
بنك بنسلفانيا للتحاليل النحوية ٢٠ ، ٦٥ ، ٦٠	الإعجام ١٠ ، ١٦ ، ٢٦ ، ٢١٤ ، ٢١٥ ، ٢١٦ ، ٢٦١
بنك كولومبيا الشجري للتحاليل النحوية الشجرية للغة العربية ١٨٣ ، ١٨٣ ، ١٨٦ ، ١٨٧ ، ١٨٨	الألفباء العربية ١٥ ، ١٧ ، ٢١٤ ، ٢٦١
	الألفون ١٣ ، ٤٨ ، ٥٢ ، ٢٩٠ ، ٧٤ ، ٧٧ ، ١٢١ ، ٢٩٠ ، ٢٩٨
	الإملاء الصرف - فونيمي ٥٢
	الإملاء المعجمي /المورفيمي ٥٩

الترجمة الآلية ٧، ٤٣، ٦٢، ١٣٧،

١٥٥، ١٦٠، ٢١٦، ٢٢٠،

٢٦٣، ٢٦٧

الترجمة الرمزية/المستندة على القواعد

٢١٢

تركيب النص ١٠٩، ١٣٦، ١٣٧

تركيب/تشبيك الحروف ١٦

الترميز ٤، ١٣، ١٥، ٣٨، ٤٥،

٤٦، ١١٩، ١٢٢، ١٣٠

تسوية الاحتمالات الهجائية ٣٨

تسوية الاحتمالات للحروف ٣٩

التشذيب ٤٣، ١١٥، ١١٦، ٢٩١،

٣٠١

التشكيل الآلي ٩، ٢٠، ٢٣، ٤٣،

٤٤

التصحيح الإملائي ٤٧، ٦٤

التصريف ٥٧، ٧٧، ٧٨، ١١٧،

١١٩، ٢٩١، ٢٩٤

تصريفية ٦٩، ٨٠، ٨٢، ٢٩١،

٣٠١

التطويل ٨٦

بنوك شجرية ١٩٠

البنوك الشجرية ٨٦، ١٦٦، ١٩٢،

١٩٦

البنوك الشجرية العربية ١٩٦

## ف

تأثير باس ٥٠

التجذيع ١١٥، ١٣٦، ٢٥٧، ٢٩١،

٣٠٥

التجريد ١٣٣، ١٣٥، ١٣٨، ١٣٩،

١٤٠، ٢٠٩

تجميع النص ٢١٦، ٢٩١، ٢٩٩

التحليل الإحصائي للغات الغنية

صرفياً ١٩٨، ٢٢٧

التحليل الصرفي ١١٤، ١١٦، ١٢٠،

١٥٣، ١٥٨، ١٦٢، ١٨٥

تحويل الصوت لنص ٦٥

الترادف ٢٠٠

الترتيب الألفبائي ١٧

الترجمة الإحصائية / المعتمدة على

المدونة ٢١٢

الجذر ١٣١، ١٣٢، ١٣٣، ١٣٤،

٢٥٤، ٢٥٥

جمع التفسير ٨٢، ٩٥، ٩٧، ٩٩،

١٠١، ١٠٢، ١١٤، ١٧٦

جمع التفسير للمؤنث ٩٧، ٩٨

الجيماتريا ١٧

### م

حرف العلة في بداية الجذع ٧٢

حركات ٢٩٣، ٣٠٤، ٣٠٥

حزمة لاتك العربية ٢٨

الحلقة ٢٧

الحوائط ٧١

### م

خريطة الأحرف ٢٨

الخط العربي ٥، ٦، ٧، ٢٠، ٢٣،

٢٥، ٣٩، ٤٠، ٥٦، ٥٩،

٢٢٧

خطوة تجهيزية ٦٤

التعابير اللغوية ١٩٩

التعرف الضوئي على الحروف ٦،

١٢، ٤١

التكتيك الصرفي ٧٧، ١٢٣، ١٥٩

تمييز الأعلام ٢٠٨

الناثر ٢١٥

التهجئة العربية ٧، ٤٧، ٥٥، ٦١

توكن ١٢٠، ١٥٢، ١٥٣، ١٥٦،

١٦١، ١٦٣

التوليد الصرفي ٧، ١١٣، ١١٤،

١٥٦، ١٦١، ٢٥٨

توليد الكلام آلياً ٤٧، ٤٧، ٦٦

توليفات الصوامت ٧٤

### ن

الشاء ٦، ٦٩، ١١٦، ١٨٢، ١٨٨،

ثنائي الإعراب ٢٩٢، ٢٩٩

### ج

جداول التوافق ١٢٣، ١٢٤، ١٢٥،

١٢٧

شبكة الكلمات العربية ٢٠٥، ٢٠٦،

٢٠٧، ٢٢٤

## ص

صامت ٥٠، ٥٢، ٧٤، ١٠٥،

١٠٧، ١١٠، ١١١

الصرف ٣، ٧، ١٦٦، ١٨٢، ١٨٥،

٣٠٠، ٣٠١، ٣٠٢، ٣٠٦

الصرف الشكلي ١١٦، ١٦٢، ١٨٥،

الصرف اللصقي ١٦٦

الصرف الوظيفي ١٨٦، ٢٩٤، ٣٠٠

الصرفية ١٠٥

الصرف - نحوي ٢١٧

صرف نحوي ٢٩٤، ٣٠٢

صفحة الشفرة ٢٨

صوامت متوافقة ١٠٧

صوائت طويلة ٥٠، ١٠٧

صوائت قصيرة ٥٠، ١٠٦، ١٠٧،

٢٩٣، ٣٠٤، ٣٠٥

الصيغة الصرفية ٧٥

صيغة الفعل ٩٣، ٢٩٤، ٣٠١

صيغي ٧٣، ١٠٥، ١٣١، ١٤٩

## ر

الرسم ٢٤٢

الرومنة ٣٦، ٤١، ٦٤

رومنة الخط العربي ٩

## ز

الزوائد ٧١، ٧٢، ٧٤، ٦٤، ٨٩،

١٠٥، ١٠٨، ١١٥، ١٣٩،

١٤٠، ١٥٦، ١٦٢، ٢٩٠،

٢٩٣، ٢٩٥

الزوائد في نهاية الكلمة ٢٩٠، ٣٠٠

## س

سلاسل المتواليات ١٢٧

سلسلة إن قرام ٤٩

السوابق ٧١، ٧٢، ٨٤، ١٥٩،

السياق ٤، ٢٩، ٤٣، ٥٨، ٦٠،

١١٨، ١٢٢، ١٣١، ١٤١،

١٥٣، ١٥٤، ٢٠٦

## ش

شبكة الكلمات الأوربية ٢٠٥، ٢٠٦

**غ**

غير صيغي ٧٣، ٧٥

**ف**

فئة السلوك المصرفي ١٣٠، ١٣١

فصحى التراث ٢

فصحى العصر ٢

فك اللبس المصرفي ١١٤، ١١٥،

١١٦، ٢٦٠

فنولوجي ٢٩٤، ٣٠٢

الفنولوجي ٣٦، ٤٧، ٤٨، ٥٠،

٥٢

الفونيم ١٣، ٤٨، ٤٩، ٥٢، ٥٩،

٦٢، ٦٦، ١٠٦

الفونيمات ١٠٦

الفيولر افريكت ٥٣

**ق**

قرافيم ١٢، ١٣، ٢٩، ٣٤، ٣٩،

٤٠

قضايا التقييس ٦٠

**ض**

الضبط بالشكل ٢٠

الضمائر المتصلة ٥٧، ٨٣، ٨٧،

١٠٩، ١١٠، ١١٧، ١٣٩،

١٥٦، ١٥٨

الضمائر الفارغة ١٩٤

**ط**

الطاء الصغيرة ٢٧

**ع**

العربية ٦٤، ٧٦، ٢٢٠

عدم التطابق المتعدد ١٠٠

عدم تطابق الجنس ٩٨

علاقة الاشتمال ٢٠٦

علم الأصوات ٦، ٤٧، ٤٨، ٤٩،

٦٦، ١٠٥، ٢٩٤، ٢٩٥،

٣٠٢، ٣٠٣

علم الأصوات العربية ٦، ٤٧، ٦٦

علم التهجئة ٥٥

علم الدلالة ١٩٩

لفظة ١، ٨٠، ٨١، ٩٤  
 اللهجات العربية ١، ٥، ٤٧، ٨٩،  
 ١٠٦، ١٢٢، ١٩٧، ١٩٧،  
 ٢٣٢، ٢٢٨  
 اللهجة الخليجية ٣  
 اللهجة الشامية ٣، ٥٤  
 لهجة شمال أفريقيا ٣  
 اللهجة اليمنية العربية ٤  
 اللواحق ٧١، ٧٢، ١٠٦، ١٢٤،  
 ١٥٩، ١٦٨  
 اللواحق ٤٣، ٧١، ٧٢، ٨٥، ٨٧،  
 ٩٢، ٩٤، ١٠٧، ١١٩، ١٢١،  
 ١٢٩، ١٣٣، ١٦٦، ٢١٤

### م

المادة المعجمية ٨٠، ٨١، ١١٥،  
 ١١٧، ١٣٤، ١٥٣، ١٥٤،  
 ١٦٢، ١٨٢، ٢٦٠  
 مبنى ٢٩٢، ٢٩٩  
 متحدة المخرج ٧٤  
 مترجم صخر ٢٢٠

القواعد المورفونيمية ١٠٥، ١٣٣

### ك

كاتب ٢٠، ٣١، ٧٣، ١٤٤،  
 ١٤٨، ١٥٦، ١٩٠، ١٩٢،  
 ٢٤٦، ٢١٧  
 كتابة صورية ٢١٤  
 الكتابة الصوتية ٢٩٥، ٣٠٥  
 كشف اللبس الصرفي ٧  
 الكشيدة ٢٥، ٨٦  
 الكلمات المقترضة ٦١

### ل

اللسانيات الحاسوبية ٢٢٣، ٢٩٦،  
 ٢٩٨  
 اللصق ٧١، ٧٧، ٧٨، ٨٣، ٨٨،  
 ١٣٢، ١٦٦  
 اللغة العربية الفصحى ٢، ٣، ٥، ٢٤،  
 اللغة العربية المعاصرة ٢، ٤، ٥، ٥٢،  
 ٥٤، ١٠٦، ١٢٨، ١٣٠،  
 ١٣١، ١٩٧

مشكلة القدس ٦٤	مجموعة الأحرف ٢٨
مشكلة القذافي ٦٢ ، ٦٤	مجموعة السمات الموسعة المقلصة ١٤٧
مشكلة حسن ٦٣	مجموعة سمات بادت ١٥١
مشكلة ١٣ ، ٢١ ، ٢٢ ، ٧٢ ، ٧٥ ، ١٣٦ ، ١٥٨ ، ١٥٩ ، ٢٠٨	مجموعة سمات يبيز ١٤٤ ، ١٤٦ ، ١٤٨
٢١٤ ، ٢١٦ ، ٢١٧ ، ٢٤٤	مجموعة سمات خوجة ١٤٩ ، ١٥٠ ، مجيد ١٢٨ ، ١٢٩ ، ١٣٠ ، ١٣٣ ، ١٣٤ ، ٢٥٩
مشكلة ماري ٦٣	المحارف ٢٨
مصنف آلة ١٥٣	المحدد ٤٨ ، ٧٤ ، ٧٩ ، ١١٩ ، ١٢٧ ، ١٥٩ ، ١٦٥ ، ١٨٢ ، ٢٠٦
المعجم ٣٨ ، ٥٩ ، ٧٨ ، ١٣٥ ، ١٤٠ ، ١٥٣ ، ١٨٢ ، ٢٠٢	المحرك التحليلي ١٢٣ ، ١٢٥ ، محلل باكوالترا الصرفي ١٢٣ ، ٢٥٨ ، محلل بايكل ٢٦١
٢٤٤ ، ٢٦٠ ، ٢٦٢	مدار ٢٦٦
معجمية ٥٧ ، ١٢٢ ، ١٢٤ ، ١٨٩ ، ٢٠٥	المدونات المتوازية ٢١٢ ، ٢١٩ ، المدونات ثنائية اللغة ٢١٢ ، المرجانة ١٢٦ ، ١٢٧ ، ١٢٨ ، ١٥٣ ، ١٥٦ ، ١٦١ ، ٢٥٩
المعدلات/المكملات الفعلية ٢١٨	مركز اتحاد البيانات اللغوية ١٨٤ ، مركز ائتلاف البيانات اللسانية ٣٦
معرفة ٦ ، ٩٢ ، ١٠٠ ، ١٤٧ ، ١٥٣ ، ١٦١ ، ١٨٠ ، ١٨١	
المعهد الوطني للمعايير والتقنية ٢٦٧	
المفردات غير المعروفة	
مُقطع النص البسيط ٦٦ ، ٦٧ ، ٢١٥	
المنوع من الصرف ١٠٢	
المُميّز ١٧٩	

- منظمة المقاييس العالمية ٣٣  
المواد المعجمية ٨١، ١١٦، ١١٧،  
١٤٠، ٢١٦
- مورفولوجي ١٦٥، ٢٤٦، ٢٥٤
- المورفيم الإعلالي ٧٦
- المورفيم السلسلي ٧١
- المورفيم الصيغي ٧١
- مورفيم ١٨، ٥٩، ٩٨، ١٠١،  
١٢٤، ١٢٧، ١٣٩، ١٤٠،  
١٤٣
- المورفيمات الصيغية ٧٣، ٧٧، ٨٢،  
١٠٥، ١٢٩
- المورفيمات ٨٢، ١٠٥، ١٣٩
- النبر ٥٣
- نظام لاتك ٢٨
- النقحرة، ٦، ٩، ٣٦، ٦٤، ٨٤،  
٢٩٢، ٣٠٢
- النقل الكتابي ٦، ٩، ١٤، ٦٤،  
٨٤، ١٢٤، ٢٩٧، ٣٠٢
- نكرة ٢٠، ٥٦، ٦٠، ٢١٩، ٣٠١
- نموذج ماركوف المخفي ٤٢  
نهاية الجذع ٥٨، ١٥٩
- هـ
- الهايرنمي ٢٠٥
- الهايوني ٢٠٥
- الهفت ٢٧
- الهومونيمي ٢٠٠
- و
- الوحدة الصرفية ٧١
- وحدة لفظية ١٢٥، ١٣٤، ١٤٠،  
١٨٨، ١٨٩، ٢٩٤، ٣٠٣
- وسط الجذع ٥٨
- الوسوم العلائقية ١٨٩
- ي
- اليونيكود ٢٩، ٣٥، ٣٦، ٣٨، ٤٠