

# Majdi Shaker Sawalha

## Personal Information

- ✚ Date / Place Of Birth: 18-Apr-1979 Irbid /Jordan
- ✚ Nationality: Jordanian
- ✚ Place of Residence: Amman, Jordan
- ✚ Postal Address: P.O.Box 3950, Irbid – 21110, Jordan
- ✚ Work Address: Computer Information Systems Department, King Abdullah II School for Information Technology, The University of Jordan, Amman – 11942, Jordan. Tel: +962 6 5355 000 ext: 22619

## Education

- ✚ **Nov 2007 to Jan 2012** **PhD, University of Leeds, Leeds, UK,**  
Natural Language Processing Group, School of Computing  
**Supervisor: Dr. Eric Atwell,** Thesis submission date: Oct-2011
- ✚ PhD Thesis Title: **Open-source Resources and Standards for Arabic Word Structure: Fine Grained Morphological Analysis of Arabic Text Corpora**  
Available on: <http://etheses.whiterose.ac.uk/2165/>
- ✚ [http://etheses.whiterose.ac.uk/2165/1/Sawalha\\_MSS\\_Computing\\_PhD\\_2011.pdf](http://etheses.whiterose.ac.uk/2165/1/Sawalha_MSS_Computing_PhD_2011.pdf)
- ✚ **Sep 2001 to Jun 2004** **M.Sc. in Computer Information Systems,** Yarmouk University, Irbid, Jordan.
- ✚ **Sep 1997 to Jun 2001** **B.Sc. in Computer Science,** Yarmouk University, Irbid, Jordan.
- ✚ **July 1997** **High School Education,** Hitteen Secondary School, Irbid, Jordan.

## Work Experiences

### **The University of Jordan, Jordan**

- ✚ Feb 2012 to date, **Assistant Professor,** Computer Information Systems Department, The University of Jordan, Amman, Jordan.
- ✚ Feb 2006 to Nov 2007, **Lecturer** – Full time, Computer Information Systems Department, The University of Jordan, Amman, Jordan.

### **The University of Leeds, UK**

- ✚ Apr 2013 to Apr 2015, **Visiting Researcher,** School of Computing, Faculty of Engineering, The University of Leeds, Leeds, UK. UK's Engineering and Physical Sciences Research Council (EPSRC) funded a 2-year project (2013-2015) of about £350K, entitled: **"Natural Language Processing Working Together with Arabic and Islamic Studies"**.

- ✚ Jun 2011 to Dec 2011, **Research Fellow**, School of Computing and School of Modern Languages and Cultures, The University of Leeds, Leeds, UK. UK's Higher Education Academy (HEA) funded a project entitled: "**A Web-as-Corpus approach to populating Wikiversity for teaching about Islam and Muslims in language, linguistics and area studies**".

### **Oman College of Management and Technology, Sultanate of Oman**

- ✚ Oct 2004 to Aug 2005, **Lecturer**, Department of Computer Science, Oman College of Management and Technology, [www.omacollege-barka.com.om](http://www.omacollege-barka.com.om), Barka, Sultanate of Oman.

### **Irbid District Electricity Company, Jordan**

- ✚ Oct 2001 to Jul 2004, **Software Developer**, Irbid District Electricity Company Ltd. (IDECO) E-mail: [ideco@go.com.jo](mailto:ideco@go.com.jo), Irbid, Jordan,

---

### **Conferences and Workshops Organization**

- ✚ **NTIT-2017: New Trends in Information Technology 2017 Conference**. 25-27 April 2017, The University of Jordan, Amman, Jordan. <http://www.ntit-conf.com>
- ✚ **NTACCL-2017: New Trends in Arabic Computational and Corpus Linguistics**, Post – NTIT-2017 Workshop. April, 27, 2017. The University of Jordan, Amman, Jordan. <http://www.ntit-conf.com/?page=new-trends-in-arabic-computational-and-corpus-linguistics-ntaccl-2017>
- ✚ **LRE-Rel 2: 2<sup>nd</sup> Workshop on Language Resource and Evaluation for Religious Texts**; LREC 2014 post-conference workshop 31st May 2014 (Afternoon session); Reykjavik, Iceland, <http://www.lrec-conf.org/proceedings/lrec2014/index.html>
- ✚ A Special Session on **Arabic Natural Language Processing: Algorithms, Resources, Tools, Techniques and Applications**, In conjunction with the First International Conference on Communications, Signal Processing, and their Applications (ICCSA'13) Sharjah, UAE, February 12 - 14, 2013 <http://www.iccsa-conf.org/>
- ✚ **LRE-Rel: Language Resource and Evaluation for Religious Texts**; LREC 2012 pre-conference workshop 22 May 2012 (Morning); Istanbul, Turkey, <http://www.comp.leeds.ac.uk/arabic/lre-rel.html>

---

### **Teaching Experience**

During my work as an assistant professor at Computer Information Systems Department, I have taught the following courses:

#### **Fall Semester 2018**

- [1] Innovative Information Systems – 3<sup>rd</sup> year undergraduate course.
- [2] Certified Software (Programming Using Python 3) – 3<sup>rd</sup> year undergraduate course.
- [3] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- [4] Basics of Computing – 1<sup>st</sup> year undergraduate course.

#### **Summer Semester 2018**

- [1] Artificial Intelligence and Expert Systems – Master's degree course.
- [2] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
- [3] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.

#### **Spring Semester 2018**

- [1] Artificial Intelligence and Expert Systems – Master's degree course.
- [2] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.

- [3] Innovative Information Systems – 3<sup>rd</sup> year undergraduate course.
- [4] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
- [5] Computational Linguistics – PhD course – Faculty of Foreign Languages

- Fall Semester 2017**
- [1] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [2] Innovative Information Systems – 3<sup>rd</sup> year undergraduate course.
  - [3] Certified Software (Programming Using Python 3) – 3<sup>rd</sup> year undergraduate course.
  - [4] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
  - [5] Computational Morphology – PhD course – Faculty of Foreign Languages.
- Summer Semester 2017:**
- [1] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2017:**
- [1] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
  - [3] Corpus Linguistics – PhD course – Faculty of Foreign Languages
- Fall Semester 2016:**
- [1] Innovative Information Systems – 3<sup>rd</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2016:**
- [1] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
  - [3] Remedial Computer Skills – 1<sup>st</sup> year undergraduate course.
- Fall Semester 2015:**
- [1] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [2] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
  - [3] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Summer Semester 2015:**
- [1] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2015:**
- [1] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [2] Certified Software (Programming Using Python 3) – 3<sup>rd</sup> year undergraduate course.
  - [3] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
  - [4] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Fall Semester 2014:**
- [1] Natural Language Processing – Masters Course.
  - [2] Artificial Intelligence – 4<sup>th</sup> year undergraduate course.
  - [3] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2014:**
- [1] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Fall Semester 2013:**
- [1] Database Management Systems – 3<sup>rd</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2013:**
- [1] Database Management Systems – 3<sup>rd</sup> year undergraduate course.
  - [2] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
  - [3] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Fall Semester 2012:**
- [1] Database Management Systems – 3<sup>rd</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Summer Semester 2012:**
- [1] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.
- Spring Semester 2012:**
- [1] Documentation and Ethics – 2<sup>nd</sup> year undergraduate course.
  - [2] Computer Skills for Humanities – 1<sup>st</sup> year undergraduate course.

## **Graduate Students Supervision**

- [1] Manal Al-Mayah, "Quranic Arabic WordNet". Department of Computer Information Systems, The University of Jordan, 2015.
- [2] Ala'a Al-Solman, "Prediction of Punctuation Marks for Classical and Modern Standard Arabic Text". Department of Computer Information Systems, The University of Jordan. 2016.
- [3] Khetam Al-Yassin, "Stochastic Approach of Part-of-Speech Tagging for Classical and MSA Arabic Text". Department of Computer Information Systems, The University of Jordan. 2017.
- [4] Nawras Al-Masri, "Spell-Checking and Correction for Arabic Text". Department of Computer Information Systems, The University of Jordan. 2017.
- [5] AbuBakr Baras, "New Arabic WordNet". Department of Computer Information Systems, The University of Jordan. 2017.
- [6] Ola Moneer, "Predicting Cancer Survival Rate: A Comparative Study", Department of Computer Information Systems, The University of Jordan. 2018.
- [7] Nazik Hassounah, "Deep Learning for Predicting Survivability for Leukemia Patients on SEER Data", Department of Computer Information Systems, The University of Jordan. 2018.
- [8] Motasem Yakhola, "Constructing an Arabic Morphological Lexicon", Department of Computer Information Systems, The University of Jordan, 2018.

## **External M.sc. Theses – Examinations Committee Membership**

- [3] Al-Khalaylah, Enas (2018) "تقييم المحللات النحوية والصرفية: دراسة في ضوء اللسانيات الحاسوبية" "Evaluating of Morphological and Syntactic Analyzers: A study in the light of Computational Linguistics", M.A thesis, Arabic Language Department, The Hashemite University.
- [2] Al-Rabab'ah, Zein A. (2016) "Thesaurus-Based Query Expansion for the Retrieval of Arabic Medical Text". M.sc. Thesis, Yarmouk University, Irbid, Jordan.
- [1] Shmaisani, Hashem A. J. (2013) "Using Semantic Approaches to Answer Questions from the Holy Quran". M.sc. Thesis, Philadelphia University, Amman, Jordan.

## **Internal M.sc Theses – Examination Committee Membership**

- [7] Al-Omari, Laith (2018) Automatic Speech Emotions Recognition for Arabic Language. M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [6] Al-latayfah, Marwa (2018) "Improving Software Defects Prediction using Feature Selection and Data Mining Techniques" M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [5] Al-Zou'bi, Mohammed (2017). Hybrid Modeling of Modern Standard and Dialectal Arabic for Automatic Continuous Speech Recognition. M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [4] Abu Sukhon, Rasha (2016) "The Design and Implementation of A Conversational Agent for Arabic Learning Environment", M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [3] Dawood, Batool (2015). "The Design and Implementation of an Arabic Question Answering System Based on the Universal Network Language (UNL)". M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [2] Shdaifat, Aala' (2015). "Towards Multithreading Sliding Window to Improve Exact Pattern Matching in Searching Text Documents". M.sc. Thesis, The University of Jordan, Amman, Jordan.
- [1] Hammad, Safia (2015). "Automatic continuous Speech Recognition Using Sub-Words Approach for Arabic Language". M.sc. Thesis, The University of Jordan, Amman, Jordan.

## **Funded Research Projects**

- ❏ Al-Shdaifat, Abdullah (PI); **Sawalha, Majdi** (Co-I); Yagi, Sane (Co-I); Al-Qudah, Mohammed (Co-I) (2017 – 2019) “Constructing the Arabic E-Dictionary” Funded by Scientific Research Fund, Ministry of Higher Education and Scientific Research in Jordan. Project’s is JOD 71,555.
- ❏ Hammo, Bassam (PI); **Sawalha, Majdi (Co-I)**; Yagi, Sane (Co-I), Al-Shdaifat, Abdullah (Co-I) (2017-2020) “Building a Syntactic Parser for Arabic”, Funded by Deanship of Academic Research, The University of Jordan. Project’s budget is JOD 12,600.
- ❏ Kedher, Mohammed Zeki (PI); Saoudi, Mohammed (Co-I); **Sawalha, Majdi (Co-I)**; Ababneh, Sami (Co-I); Hamdan, Yousef (Co-I); Hattab, Mamoun (Co-I) (2017-2018) “Index for Arabic Natural Language Processing Research”, Funded by Arabic Language Academy in Jordan. Project’s budget is JOD 35,000.
- ❏ Awad, Zahra (PI); **Sawalha, Majdi (Co-I)** and Kirner-Ludwig, Monika (Co-I) (2016-2019) “Arab Women in Western Press During Arab Spring: Image and Perception in Germany”, Funded by Deanship of Academic Research, The University of Jordan. Project’s budget is JOD 9,400.
- ❏ Atwell, Eric (PI); Dickins, James (Co-I); Brierley, Claire (Researcher) and **Sawalha, Majdi (Researcher)** (2013-2015) “Natural Language Processing Working Together with Arabic and Islamic Studies” Funded by UK’s Engineering and Physical Sciences Research Council (EPSRC). Project’s Budget is £350,000.
- ❏ Atwell, Eric (PI); Dickins, James (Co-I) and **Sawalha, Majdi (Researcher)** (2011) “A Web-as-Corpus approach to populating Wikiversity for teaching about Islam and Muslims in language, linguistics and area studies”. Funded by UK’s Higher Education Academy (HEA). Project’s Budget is £3,000.

## Publications

The full list of publications is found on the following websites:

**Google scholar:** <http://scholar.google.com/citations?user=6LD7HyIAAAAJ&hl=en>  
**Research Gate:** [https://www.researchgate.net/profile/Majdi\\_Sawalha](https://www.researchgate.net/profile/Majdi_Sawalha)  
**ORCID ID:** 0000-0003-4975-6709  
**ResearcherID:** [www.researcherid.com](http://www.researcherid.com) - E-1107-2016  
**Website:** <http://eacademic.ju.edu.jo/sawalha.majdi/default.aspx>

## PhD Thesis

Sawalha, MSS (2011). **Open-source resources and standards for Arabic word structure analysis: Fine grained morphological analysis of Arabic text corpora**, PhD thesis, School of Computing, University of Leeds.

## Conference / Workshop Proceedings Editing

- [1] **Sawalha, Majdi** (2017) New Trends in Information Technology (NTIT-2017) International Conference. 25-27 April 2017, The University of Jordan, Amman, Jordan, <http://www.ntit-conf.com>
- [2] Brierley, Claire; **Sawalha, Majdi**; Atwell, Eric (2014) **LRE-Rel2: 2<sup>nd</sup> Language Resources and Evaluation for Religious Texts**. LREC 2012 pre-conference workshop 22 May 2012 (Morning); Istanbul, Turkey, <http://www.comp.leeds.ac.uk/eric/lre-rel/lre-relProceedings.pdf>
- [3] Atwell, Eric; Brierley, Claire; **Sawalha, Majdi** (2012) **LRE-Rel: Language Resources and Evaluation for Religious Texts**. LREC 2012 pre-conference workshop 22 May 2012 (Morning); Istanbul, Turkey, <http://www.comp.leeds.ac.uk/eric/lre-rel/lre-relProceedings.pdf>

## Journal papers

- [12] **Sawalha, Majdi** (2019) "The Design and Construction of the Traditional Arabic Lexicons Corpus (The TAL-Corpus)", **Modern Applied Science**. Vol. 13, No. 2 (2019) DOI:10.5539/mas.v13n2p95
- [11] Mustafa-Awad, Zahra; **Sawalha, Majdi**; Kerner-Ludwig, Monika and Tabaza, Dua'a (2019-Forthcoming) "Arab women in Western press: Designing news corpora for Arab women in British, American, and German news media during the Arab Spring", **Research in Computing Science**.
- [10] Brierley, Claire; **Sawalha, Majdi**; Islam, Tajul; Dickins, James; Atwell, Eric (2018) "Automatic Extraction of Quranic Lexis Representing Two Different Notions of Linguistic Saliency: Keyness and Prosodic Prominence", **Journal of Semitic Studies**, Volume 63, Issue 2, 1 October 2018, Pages 407–456, <https://doi.org/10.1093/jss/fgy005>
- [09] **Sawalha, Majdi**; Brierley, Claire, Atwell, Eric and Dickins, James (2017), "Text Analytics and Transcription Technology for Quranic Arabic", **International Journal on Islamic Applications in Computer Science And Technology**; Vol 5, No 2.
- [08] Hussein, Riyadh; **Sawalha, Majdi** (2016) *A Corpus-based Study of Similes in British and American English*. **Arab World English Journal**, Volume 7, Issue 2, June 2016. Pp. 49-60 DOI: <https://dx.doi.org/10.24093/awej/vol7no2.4>
- [07] Brierley, Claire; **Sawalha, Majdi**; Heselwood, Barry; Atwell, Eric. (2016) *A Verified Arabic-IPA Mapping for Arabic Transcription Technology, Informed by Quranic Recitation, Traditional Arabic Linguistics, and Modern Phonetics*. **Journal of Semitic Studies**, (Spring 2016) 61 (1): 157-186. doi: 10.1093/jss/fgv035. Oxford Journals.

- [06] AlMaayah, Manal; **Sawalha, Majdi**; Abushariah, Mohammed (2016) *Towards an automatic extraction of synonyms for Quranic Arabic WordNet*. **International Journal of Speech Technology**, Volume 19, Issue 2, pp 177–189, June 2016, DOI: 10.1007/s10772-015-9301-9
- [05] Abu Shawar, Bayan; Atwell, Eric; **Sawalha, Majdi** (2014) *A Web-as-Corpus approach to populating Wikiversity for teaching information technology modules*. **International Journal of Advances in Electronics Engineering – IJAEE Volume 4, issue 1, Pp. 28-32**.
- [04] **Sawalha, Majdi**; Atwell, Eric (2013). *A standard tag set expounding traditional morphological features for Arabic language part-of-speech tagging*, **Word Structure** 6 (1), 43-99, Edinburgh University Press.
- [03] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric (2012). *Prosody Prediction for Arabic via the Open-Source Boundary-Annotated Qur'an Corpus*, **Journal of Speech Sciences** 2 (2), 175-191.
- [02] Kanaan, Ghassan; Hammouri; Awni, Al-Shalabi, Riyad; **Sawalha, Majdi** (2009). *A New Question Answering System for the Arabic Language*. **American Journal of Applied Sciences** 6 (4): 797-805, 2009
- [01] Kanaan, Ghassan., Al-Shalabi, Riyad; **Sawalha, Majdi** (2005). *Improving Arabic Information Retrieval Systems Using Part of Speech Tagging*. **Information Technology Journal** 4(1): 32-37. 2005

### **Book Chapters**

- [1] Brierley, Claire; **Sawalha, Majdi**; El-Farahaty, Hanem (2019-Forthcoming) *Translating sacred sounds: encoding tajwīd rules in automatically-generated IPA transcriptions of Quranic Arabic*. Routledge Handbook of Arabic Translation. Hanna, Sameh F.; El-Farahaty, Hanem and Khalifa, A. W. London/New York, Routledge.
- [2] Mustafa-Awad, Zahra; Kerner-Ludwig, Monika and **Sawalha, Majdi** (2019-Forthcoming) *Arab women in Western news discourse during the Arab Spring: Fading stereotypes and emerging images*, in Garzone, Giuliana; Logaldo, Mara; Santuili, Francesca (Eds) Language of Conflicts: Polarisation/Popularisation of Discourse in the Periodical Press. Linguistic Insights – Studies in Language & Communication Series, Peter Lang International Academic Publisher.

### **Conference proceedings**

- [30] Abu Elberak, Ola; Alnemer, Loay; **Sawalha, Majdi**; and Alsakran, Jamal (2019-forthcoming) "Predicting Cancer Survivability: a Comparative Study", International Conference on Emerging Internet, Data and Web Technologies (EIDWT-2019), February 26 to February 28, 2019, Higher Colleges of Technology, Fujairah Campus, United Arab Emirates (UAE).
- [29] Mustafa-Awad, Zahra; Kerner-Ludwig, Monika and **Sawalha, Majdi** (2017) "Arab women in news discourse during the Arab Spring: Broken stereotypes and emerging images", Conflict in the Periodical Press: 6<sup>th</sup> International Conference of the European Society for Periodical Research (ESPRit), IULM (International University of Languages and Media), Milan, Italy. 28-30 June 2017
- [28] Sawalha, Noor and **Sawalha, Majdi** (2017) "A study of Arabic Keyboard Layout", in Proceedings of NTIT-2017: New Trends in Information Technology, The University of Jordan, Amman, Jordan, 25 - 27 April, 2017
- [27] Yassen, Khetam; **Sawalha, Majdi** and Alzaghoul, Fawaz (2017) "Part-of-Speech Tagging for Classical and MSA Arabic Text Using NLTK", in Proceedings of NTIT-2017: New Trends in Information Technology, The University of Jordan, Amman, Jordan, 25 - 27 April, 2017.
- [26] Mustafa-Awad, Zahra; **Sawalha, Majdi**; Kerner-Ludwig, Monika and Tabaza, Dua'a (2017) "Arab women in Western press: Designing news corpora for Arab women in British, American, and German news

media during the Arab Spring” In: proceedings of The 18<sup>th</sup> International Conference on Computing and Computational Linguistics CICLING 2017, Budapest (Hungary), 17-23 April 2017.

- [25] Mustafa-Awad, Zahra; Kerner-Ludwig, Monika; and **Sawalha, Majdi** (2016) “Arab Women in Western Print Press during the Arab Spring: Image and Perception in Germany” In: proceedings of The International Conference Language, Literature and Culture in Education 2016 (LLCE2016), Venice (Italy), 12 – 14 July 2016
- [24] Al-Solman, Ala’a; **Sawalha, Majdi** and Yagi, Sane (2015) “Punctuation Marks in Arabic: A Computational Linguistic Approach”. In: proceedings of WCIS 2015: Arabic Natural Language Processing: Systems, Models and Applications, 2<sup>th</sup> – 13<sup>th</sup> October 2015, Tabuk, KSA.
- [23] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric and Dickins, James (2014) “Text Analytics and Transcription Technology for Quranic Arabic”, in: proceedings of IMAN 2014: 2nd International Conference on Islamic Applications in Computer Science and Technologies, 12<sup>th</sup> – 13<sup>th</sup> October 2014, Amman, Jordan
- [22] **Sawalha, Majdi**; Al-Humssi, Laila; Momani, Raya; AlSaber, Sereen (2014). Automatic Qur’an Reciter. in Proceedings of the 2<sup>nd</sup> International Conference on Islamic Applications in Computer Science and Technologies (IMAN 2014), 12<sup>th</sup> – 13<sup>th</sup> October 2014, Amman, Jordan.
- [21] **Sawalha, Majdi**; Brierley, Claire and Atwell, Eric (2014) "Automatically generated, phonemic Arabic-IPA pronunciation tiers for the Boundary Annotated Qur'an Dataset for Machine Learning (version 2.0)", in: **proceedings of LRE-Rel 2: 2nd Workshop on Language Resource and Evaluation for Religious Texts, LREC 2014 post-conference workshop** 31st May 2014, Reykjavik, Iceland
- [20] Brierley, Claire; **Sawalha, Majdi**; and Atwell, Eric (2014) Tools for Arabic Natural Language Processing: a case study in qalqalah prosody, in: **proceedings of LREC 2014 the 9th edition of the Language Resources and Evaluation Conference**, 26-31 May, Reykjavik, Iceland
- [19] AlMaayah, Manal; **Sawalha, Majdi** and Abushariah, Mohammad A. M. (2014). A Proposed Model for Quranic Arabic WordNet, in: **proceedings of LRE-Rel 2: 2nd Workshop on Language Resource and Evaluation for Religious Texts, LREC 2014 post-conference workshop** 31st May 2014, Reykjavik, Iceland
- [18] **Sawalha, Majdi**; Atwell, Eric (2013). Comparing morphological tag-sets for Arabic and English. **Proceedings of the 7th International Corpus Linguistics Conference CL2013**, 22-26 July 2013, Lancaster, UK.
- [17] **Sawalha, Majdi**; Atwell, Eric; Abushariah, Mohammad A. M. (2013). Accelerating the processing of large corpora: using Grid Computing for lemmatizing the 176 million words Arabic Internet Corpus. **Proceedings of the 2nd Workshop of Arabic Corpus Linguistics WACL-2**, 22 July 2013, Lancaster, UK.
- [16] Abbas, Noorhan; AlDhubayi, Luluh; Al-Khalifa, Hend; Alqassem, Zainab; Atwell, Eric; Dukes, Kais; **Sawalha, Majdi**; Sharaf, Abdul-Baquee Mohammed (2013). Unifying linguistic annotations and ontologies for the Arabic Quran. **Proceedings of the 2nd Workshop of Arabic Corpus Linguistics WACL-2**, 22 July 2013, Lancaster, UK.
- [15] Abushariah, Mohammad A. M; **Sawalha, Majdi**; (2013). The effects of speakers' gender, age, and region on overall performance of Arabic automatic speech recognition systems using the phonetically rich and balanced Modern Standard Arabic speech corpus. **Proceedings of the 2nd Workshop of Arabic Corpus Linguistics WACL-2**, 22 July 2013, Lancaster, UK.
- [14] **Sawalha, Majdi**; Atwell, Eric; Abushariah, Mohammad A. M. (2013). SALMA: Standard Arabic Language Morphological Analysis. **Proceedings of the First International Conference on Communications, Signal Processing, and their Applications (ICCSPA'13)**. Sharjah, UAE.



- [13] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric (2012). التحليل الآلي للوقف والابتداء في نصوص اللغة العربية الحديثة. "Automatic Analysis of Phrase-Break Prediction for Arabic". **Proceedings of the International Computing Conference in Arabic: ICCA 2012**, Cairo, 26-28 Dec 2012, <http://www.altec-center.org/ICCA/11-3-4/>.
- [12] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric (2012). Predicting Phrase Breaks in Classical and Modern Standard Arabic Text. in: **Proceedings of the Language Resource and Evaluation Conference LREC 2012**, 17-23 May 2012, Istanbul, Turkey.
- [11] Brierley, Claire; **Sawalha, Majdi**; Atwell, Eric (2012). Open-Source Boundary-Annotated Corpus for Arabic Speech and Language Processing. in: **Proceedings of the Language Resource and Evaluation Conference LREC 2012**, 17-23 May 2012, Istanbul, Turkey.
- [10] **Sawalha, Majdi**; Atwell, Eric (2011). التحليل الصرفي لنصوص اللغة العربية الحديثة والكلاسيكية. "Morphological Analysis of Classical and Modern Standard Arabic Text". **7th International Computing Conference in Arabic (ICCA11)**. 31st May - 2nd June 2011, Imam Mohammed Ibn Saud University, Riyadh, KSA.
- [09] Atwell, Eric; Brierley, Claire; Dukes, Kais; **Sawalha, Majdi**; Sharaf, Abdul-Baqee (2011) An Artificial Intelligence Approach to Arabic and Islamic Content on the Internet. in: **Proceedings of NITS 3rd National Information Technology Symposium**.
- [08] Sharaf, Abdul-Baqee; Atwell, Eric; Dukes Kais; **Sawalha, Majdi**; Al-Saif, Amal; Sharoff, Serge; Markert, Katja; Al-Sulaiti, Latifa; Abu Shawar, Bayan; Abbas, Nora; Roberts, Andy (2010). المشاريع الحاسوبية على اللغة العربية والقرآن بجامعة ليدز "Arabic and Quranic Computational Linguistics Projects at the University of Leeds". in: **Proceedings of the workshop of Increasing Arabic Contents on the Web, organized by Arab League Educational, Cultural and Scientific Organization (ALECSO)**, 16-19 October 2010, Damascus, Syria.
- [07] **Sawalha, Majdi**; Atwell, Eric (2010). *Fine-Grain Morphological Analyzer and Part-of-Speech Tagger for Arabic Text*. in: **Proceedings of the Language Resource and Evaluation Conference LREC 2010**, 17-23 May 2010, Valletta, Malta.
- [06] **Sawalha, Majdi**; Atwell, Eric (2010). *Constructing and Using Broad-Coverage Lexical Resource for Enhancing Morphological Analysis of Arabic*. in: **Proceedings of the Language Resource and Evaluation Conference LREC 2010**, 17-23 May 2010, Valletta, Malta.
- [05] **Sawalha, Majdi**; Atwell, Eric (2009). *Linguistically Informed and Corpus Informed Morphological Analysis of Arabic*. in: **Proceedings of the 5th International Corpus Linguistics Conference CL2009**, 20-23 July 2009, Liverpool, UK.
- [04] **Sawalha, Majdi**; Atwell, Eric. (2009) *توظيف قواعد النحو والصرف في بناء محلل صرفي للغة العربية (Adapting Language Grammar Rules for Building Morphological Analyzer for Arabic Language)*. in: **Proceedings of the workshop of morphological analyzer's experts for Arabic language**, organized by Arab League Educational, Cultural and Scientific Organization (ALECSO), King Abdul-Aziz City of Technology ( KACT) and Arabic language Academy. Damascus, Syria. 26-28 April 2009.
- [03] **Sawalha, Majdi**; Atwell, Eric. (2008) *Comparative evaluation of Arabic language morphological analysers and stemmers* in: **Proceedings of COLING 2008 22nd International Conference on Computational Linguistics**. 18-22 August, 2008, Manchester, UK.
- [02] Kanaan, Ghassan; Al-Shalabi, Riyad; **Sawalha, Majdi** (2004), Improving Arabic Information Systems using Part of Speech Tagging, in: Proceedings of SJICM, The Sixth Jordanian International Congress of Mathematics, Organized by the Jordanian Mathematical Society in cooperation with Yarmouk University, Irbid, Jordan. 31st August – 3rd September 2004

- [01] Kanaan, Ghassan; Al-Shalabi, Riyad; **Sawalha, Majdi** (2003), Full Automatic Arabic Text Tagging System, in: **Proceedings of ICITNS, International Conference on Information Technology and Natural Sciences**, Al-Zaytoonah University of Jordan, Amman, Jordan, October 19-21, 2003.

### **Conference presentations**

- [13] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric; Dickins, James (2015). *Arabic > IPA Transcription Technology*. Language Data Analytics Workshop. 16 September 2015, The University of Leeds, Leeds, UK.
- [12] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric; Dickins, James (2015). *Language Resources, Algorithms, and Machine Learning for Prosodic Analysis of Classical and Modern Standard Arabic: an Overview*. The Forum for Arabic Linguistics conference. 28-30th July 2015, the University of Essex, Essex, U.K.
- [11] Brierley, Claire; **Sawalha, Majdi**; Mansour, M. Khalid; Watson, Janet (2014). *Sharing concepts in Qur'anic recitation and modern Phonetics*. The 2014 Seminar for Arabian Studies, Friday 25<sup>th</sup> to Sunday 27<sup>th</sup> July, Clore Centre, British Museum, London, U.K.
- [10] Brierley, Claire; **Sawalha, Majdi**; Atwell, Eric; Dickins, James (2014). *Quran Recitation, Natural Language Processing, and Arabic and Islamic Studies*. Multimodality in Language Research Conference, 26 June, 2014. Leeds, U.K.
- [09] Brierley, Claire; **Sawalha, Majdi** (2014). "A Grapheme-Phoneme Mapping for Arabic Natural Language Processing", **Power To Connect In A Changing World: International Conference on Language, Literature and Translation**, 22-24 April 2014, The Faculty of Foreign Languages, The University of Jordan, Amman, Jordan.
- [08] **Sawalha, Majdi**; Atwell, Eric (2012). *Visualisation of Arabic Morphology*. **Advances in Visual Methods for Linguistics AVML 2012**. 7-9 Sep 2012, University of York, York, UK.
- [07] Brierley, Claire; **Sawalha, Majdi** (2012). *Visualisation of Prosody in English and Arabic Speech Corpora*. **Advances in Visual Methods for Linguistics AVML 2012**. 7-9 Sep 2012, University of York, York, UK.
- [06] **Sawalha, Majdi**; Brierley, Claire; Atwell, Eric (2012). *Modern Standard Arabic sentences: problems and pitfalls for automated processing*. **The 6th Inter-Varietal Applied Corpus Studies (IVACS) conference**, 21-22 June, 2012, Leeds Metropolitan University, Leeds, UK.
- [05] Brierley, Claire; **Sawalha, Majdi**; Atwell, Eric (2012). *Boundary Annotated Qur'an Corpus for Arabic Phrase Break Prediction*. The 2012 IVACS Annual Symposium, University of Cambridge, Cambridge, UK, 13th January, 2012.
- [04] Brierley, Claire; **Sawalha, Majdi**; Atwell, Eric (2011). *Arabic Phonetics and Phonology for Text Analytics and Natural Language Processing Applications*. **Postgraduate Conference on Arabic Phonetics and Phonology, PCAPP 2011**. 10th June 2011, University of York, York, UK.
- [03] **Sawalha, Majdi**; Atwell, Eric (2011). *Corpus Linguistics Resources and Tools for Arabic Lexicography*. **Workshop on Arabic Corpus Linguistics**, 11-12 April 2011, Lancaster University, Lancaster, UK.
- [02] **Sawalha, Majdi**; Atwell, Eric (2011). *Accelerating the Processing of Large Corpora: Using Grid Computing Technologies for Lemmatizing 176 Million Words Arabic Internet Corpus*. **Advanced Research Computing Open Event**, March 28th 2011, University of Leeds, Leeds, UK.
- [01] Atwell, Eric; Abbas, Noorhan; Abu Shawar, Bayan; Alsaif, Amal; Al-Sulaiti, Latifa; Roberts, Andrew; **Sawalha, Majdi** (2008). *Mapping Middle Eastern and North African diasporas: Arabic corpus linguistics research at the University of Leeds*. **BRISMES Conference 2008**. University of Leeds, Leeds, UK.

## **References**

- Dr. Eric Atwell - Eric Atwell, Associate Professor, Computing and Language; MSc Tutor; I-AIBS Institute for Artificial Intelligence and Biological Systems, School of Computing, Faculty of Engineering, UNIVERSITY OF LEEDS, Leeds LS2 9JT, England. TEL: +44-113-3435430 FAX: +44-113-3435468, WWW: <http://www.comp.leeds.ac.uk/eric> <http://www.comp.leeds.ac.uk/arabic> E-mail: [E.S.Atwell@leeds.ac.uk](mailto:E.S.Atwell@leeds.ac.uk)
- Prof. James Dickins - Professor of Arabic; Deputy Head of School (Arabic, Middle Eastern and East Asian Studies), UNIVERSITY OF LEEDS, Leeds LS2 9JT, England. E-mail: [j.dickins@leeds.ac.uk](mailto:j.dickins@leeds.ac.uk)
- Dr. Claire Brierley, Senior Research Fellow, I-AIBS Institute for Artificial Intelligence and Biological Systems, School of Computing, Faculty of Engineering, UNIVERSITY OF LEEDS, Leeds LS2 9JT, England. E-mail: [C.Brierley@leeds.ac.uk](mailto:C.Brierley@leeds.ac.uk)
- Prof. Sane Yagi – Former Dean of the Faculty of Foreign Languages at The University of Jordan. E-mail: [saneyagi@hotmail.com](mailto:saneyagi@hotmail.com)
- Prof. Zahra Mustafa Awad – Former Dean of the Faculty of Foreign Languages at The University of Jordan. E-mail: [z.awad@ju.edu.jo](mailto:z.awad@ju.edu.jo)
- Prof. Mohammad Khalid Mansour – Former Deputy Dean of the Faculty of Islamic Studies at The University of Jordan. E-mail: [drmkhm@hotmail.com](mailto:drmkhm@hotmail.com)
- H.E. Prof. Azzam Sleit– Former Dean of King Abdullah II School for Information Technology (KASIT) at The University of Jordan. E-mail: [azzam.sleit@ju.edu.jo](mailto:azzam.sleit@ju.edu.jo)
- Prof. Nadim Obaid – Former Dean of King Abdullah II School for Information Technology (KASIT) at The University of Jordan. E-mail: [nadim@ju.edu.jo](mailto:nadim@ju.edu.jo)
- Prof. Fawaz Al-Zaghoul – Former Dean of King Abdullah II School for Information Technology (KASIT) at The University of Jordan. E-mail: [fawaz@ju.edu.jo](mailto:fawaz@ju.edu.jo)