Second Workshop on Arabic Natural Language Processing

(ANLP 2015)

Co-located with ACL-IJCNLP 2015

Beijing, China
30 July 2015

ISBN: 978-1-5108-0948-2

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571

Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2015) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2015)

For permission requests, please contact the Association for Computational Linguistics
at the address below.

Association for Computational Linguistics
209 N. Eighth Street
Stroudsburg, Pennsylvania 18360

Phone: 1-570-476-8006
Fax: 1-570-476-0860

acl@aclweb.org

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

Classifying Arab Names Geographically
Hamdy Mubarak and Kareem Darwish 1

Deep Learning Models for Sentiment Analysis in Arabic
Ahmad Al Sallab, Hazem Hajj, Gilbert Badaro, Ramy Baly, Wassim El Hajj and Khaled Bashir Shaban 9

A Light Lexicon-based Mobile Application for Sentiment Mining of Arabic Tweets
Gilbert Badaro, Ramy Baly, Rana Akel, Linda Fayad, Jeffrey Khairallah, Hazem Hajj, Khaled Shaban and Wassim El-Hajj 18

The Second QALB Shared Task on Automatic Text Correction for Arabic
Alla Rozovskaya, Houda Bouamor, Nizar Habash, Wajdi Zaghouani, Ossama Obeid and Behrang Mohit 26

Natural Language Processing for Dialectical Arabic: A Survey
Abdulhadi Shoufan and Sumaya Alameri 36

DIWAN: A Dialectal Word Annotation Tool for Arabic
Faisal Al-Shargi and Owen Rambow 49

POS-tagging of Tunisian Dialect Using Standard Arabic Resources and Tools
Ahmed Hamdi, Alexis Nasr, Nizar Habash and Nuria Gala 59

A Conventional Orthography for Algerian Arabic
Houda Saadane and Nizar Habash 69

A Pilot Study on Arabic Multi-Genre Corpus Diacritization
Houda Bouamor, Wajdi Zaghouani, Mona Diab, Ossama Obeid, Kemal Oflazer, Mahmoud Ghoneim and Abdelati Hawwari 80

Annotating Targets of Opinions in Arabic using Crowdsourcing
Noura Farra, Kathy McKeown and Nizar Habash 89

Best Practices for Crowdsourcing Dialectal Arabic Speech Transcription
Samantha Wray, Hamdy Mubarak and Ahmed Ali 99

Joint Arabic Segmentation and Part-Of-Speech Tagging
Shabib AlGahtani and John McNaught 108

Multi-Reference Evaluation for Dialectal Speech Recognition System: A Study for Egyptian ASR
Ahmed Ali, Walid Magdy and Steve Renals 118

Arib@QALB-2015 Shared Task: A Hybrid Cascade Model for Arabic Spelling Error Detection and Correction
Nouf AlShenaifi, Rehab AlNefie, Maha Al-Yahya and Hend Al-Khalifa 127

CUFE@QALB-2015 Shared Task: Arabic Error Correction System
Michael Nawar 133

GWU-HASP-2015@QALB-2015 Shared Task: Priming Spelling Candidates with Probability
Mohammed Attia, Mohamed Al-Badrashiny and Mona Diab 138

QCMUQ@QALB-2015 Shared Task: Combining Character level MT and Error-tolerant Finite-State Recognition for Arabic Spelling Correction
Houda Bouamor, Hassan Sajjad, Nadir Durrani and Kemal Oflazer 144

QCRI@QALB-2015 Shared Task: Correction of Arabic Text for Native and Non-Native Speakers' Errors
Hamdy Mubarak, Kareem Darwish and Ahmed Abdelali 150

SAHSOH@QALB-2015 Shared Task: A Rule-Based Correction Method of Common Arabic Native and Non-Native Speakers' Errors
Wajdi Zaghouani, Taha Zerrouki and Amar Balla 155

TECHLIMED@QALB-Shared Task 2015: a hybrid Arabic Error Correction System
Djamel MOSTEFA, Jaber ABUALASAL, Omar ASBAYOU, Mahmoud GZAWI and Ramzi Abbès 161

UMMU@QALB-2015 Shared Task: Character and Word level SMT pipeline for Automatic Error Correction of Arabic Text
Fethi Bougares and Houda Bouamor 166

Robust Part-of-speech Tagging of Arabic Text
Hanan Aldarmaki and Mona Diab 173

Answer Selection in Arabic Community Question Answering: A Feature-Rich Approach
Yonatan Belinkov, Alberto Barrón-Cedeño and Hamdy Mubarak 183

EDRAK: Entity-Centric Data Resource for Arabic Knowledge
Mohamed H. Gad-elrab, Mohamed Amir Yosef and Gerhard Weikum 191