

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT 2019)

Long and Short Papers

Minneapolis, Minnesota, USA
2 – 7 June 2019

Volume 1 of 6

ISBN: 978-1-5108-8754-1

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© (2019) by the Association for Computational Linguistics
All rights reserved.

Printed by Curran Associates, Inc. (2019)

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: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Table of Contents

<i>Entity Recognition at First Sight: Improving NER with Eye Movement Information</i> Nora Hollenstein and Ce Zhang	1
<i>The emergence of number and syntax units in LSTM language models</i> Yair Lakretz, Germán Kruszewski, Théo Desbordes, Dieuwke Hupkes, Stanislas Dehaene and Marco Baroni	11
<i>Neural Self-Training through Spaced Repetition</i> Hadi Amiri	21
<i>Neural language models as psycholinguistic subjects: Representations of syntactic state</i> Richard Futrell, Ethan Wilcox, Takashi Morita, Peng Qian, Miguel Ballesteros and Roger Levy	32
<i>Understanding language-elicited EEG data by predicting it from a fine-tuned language model</i> Dan Schwartz and Tom Mitchell	43
<i>Pre-training on high-resource speech recognition improves low-resource speech-to-text translation</i> Sameer Bansal, Herman Kamper, Karen Livescu, Adam Lopez and Sharon Goldwater	58
<i>Measuring the perceptual availability of phonological features during language acquisition using unsupervised binary stochastic autoencoders</i> Cory Shain and Micha Elsner	69
<i>Giving Attention to the Unexpected: Using Prosody Innovations in Disfluency Detection</i> Vicky Zayats and Mari Ostendorf	86
<i>Massively Multilingual Adversarial Speech Recognition</i> Oliver Adams, Matthew Wiesner, Shinji Watanabe and David Yarowsky	96
<i>Lost in Interpretation: Predicting Untranslated Terminology in Simultaneous Interpretation</i> Nikolai Vogler, Craig Stewart and Graham Neubig	109
<i>AudioCaps: Generating Captions for Audios in The Wild</i> Chris Dongjoo Kim, Byeongchang Kim, Hyunmin Lee and Gunhee Kim	119
<i>"President Vows to Cut <Taxes> Hair": Dataset and Analysis of Creative Text Editing for Humorous Headlines</i> Nabil Hossain, John Krumm and Michael Gamon	133
<i>Answer-based Adversarial Training for Generating Clarification Questions</i> Sudha Rao and Hal Daumé III	143
<i>Improving Grammatical Error Correction via Pre-Training a Copy-Augmented Architecture with Unlabeled Data</i> Wei Zhao, Liang Wang, Kewei Shen, Ruoyu Jia and Jingming Liu	156
<i>Topic-Guided Variational Auto-Encoder for Text Generation</i> Wenlin Wang, Zhe Gan, Hongteng Xu, Ruiyi Zhang, Guoyin Wang, Dinghan Shen, Changyou Chen and Lawrence Carin	166
<i>Implementation of a Chomsky-Schützenberger n-best parser for weighted multiple context-free grammars</i> Thomas Ruprecht and Tobias Denkinger	178

<i>Phylogenic Multi-Lingual Dependency Parsing</i> Mathieu Dehouck and Pascal Denis	192
<i>Discontinuous Constituency Parsing with a Stack-Free Transition System and a Dynamic Oracle</i> Maximin Coavoux and Shay B. Cohen	204
<i>How Bad are PoS Tagger in Cross-Corpora Settings? Evaluating Annotation Divergence in the UD Project.</i> Guillaume Wisniewski and François Yvon	218
<i>CCG Parsing Algorithm with Incremental Tree Rotation</i> Miloš Stanojević and Mark Steedman	228
<i>Cyclical Annealing Schedule: A Simple Approach to Mitigating KL Vanishing</i> Hao Fu, Chunyuan Li, Xiaodong Liu, Jianfeng Gao, Asli Celikyilmaz and Lawrence Carin . . .	240
<i>Recurrent models and lower bounds for projective syntactic decoding</i> Natalie Schluter	251
<i>Evaluating Composition Models for Verb Phrase Elliptical Sentence Embeddings</i> Gijs Wijnholds and Mehrnoosh Sadrzadeh	261
<i>Neural Finite-State Transducers: Beyond Rational Relations</i> Chu-Cheng Lin, Hao Zhu, Matthew R. Gormley and Jason Eisner	272
<i>Riemannian Normalizing Flow on Variational Wasserstein Autoencoder for Text Modeling</i> Prince Zizhuang Wang and William Yang Wang	284
<i>A Study of Incorrect Paraphrases in Crowdsourced User Utterances</i> Mohammad-Ali Yaghoub-Zadeh-Fard, Boualem Benatallah, Moshe Chai Barukh and Shayan Zamani-rad	295
<i>ComQA: A Community-sourced Dataset for Complex Factoid Question Answering with Paraphrase Clusters</i> Abdalghani Abujabal, Rishiraj Saha Roy, Mohamed Yahya and Gerhard Weikum	307
<i>FreebaseQA: A New Factoid QA Data Set Matching Trivia-Style Question-Answer Pairs with Freebase</i> Kelvin Jiang, Dekun Wu and Hui Jiang	318
<i>Simple Question Answering with Subgraph Ranking and Joint-Scoring</i> Wenbo Zhao, Tagyoung Chung, Anuj Goyal and Angeliki Metallinou	324
<i>Learning to Attend On Essential Terms: An Enhanced Retriever-Reader Model for Open-domain Question Answering</i> Jianmo Ni, Chenguang Zhu, Weizhu Chen and Julian McAuley	335
<i>UHop: An Unrestricted-Hop Relation Extraction Framework for Knowledge-Based Question Answering</i> Zi-Yuan Chen, Chih-Hung Chang, Yi-Pei Chen, Jijnasa Nayak and Lun-Wei Ku	345
<i>BAG: Bi-directional Attention Entity Graph Convolutional Network for Multi-hop Reasoning Question Answering</i> Yu Cao, Meng Fang and Dacheng Tao	357
<i>Vector of Locally-Aggregated Word Embeddings (VLAWE): A Novel Document-level Representation</i> Radu Tudor Ionescu and Andrei Butnaru	363

<i>Multi-task Learning for Multi-modal Emotion Recognition and Sentiment Analysis</i> Md Shad Akhtar, Dushyant Chauhan, Deepanway Ghosal, Soujanya Poria, Asif Ekbal and Pushpak Bhattacharyya	370
<i>Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence</i> Chi Sun, Luyao Huang and Xipeng Qiu	380
<i>A Variational Approach to Weakly Supervised Document-Level Multi-Aspect Sentiment Classification</i> Ziqian Zeng, Wenxuan Zhou, Xin Liu and Yangqiu Song	386
<i>HiGRU: Hierarchical Gated Recurrent Units for Utterance-Level Emotion Recognition</i> Wenxiang Jiao, Haiqin Yang, Irwin King and Michael R. Lyu	397
<i>Learning Interpretable Negation Rules via Weak Supervision at Document Level: A Reinforcement Learning Approach</i> Nicolas Pröllochs, Stefan Feuerriegel and Dirk Neumann	407
<i>Simplified Neural Unsupervised Domain Adaptation</i> Timothy Miller	414
<i>Learning Bilingual Sentiment-Specific Word Embeddings without Cross-lingual Supervision</i> Yanlin Feng and Xiaojun Wan	420
<i>ReWE: Regressing Word Embeddings for Regularization of Neural Machine Translation Systems</i> Inigo Jauregi Unanue, Ehsan Zare Borzeshi, Nazanin Esmaili and Massimo Piccardi	430
<i>Lost in Machine Translation: A Method to Reduce Meaning Loss</i> Reuben Cohn-Gordon and Noah Goodman	437
<i>Bi-Directional Differentiable Input Reconstruction for Low-Resource Neural Machine Translation</i> Xing Niu, Weijia Xu and Marine Carpuat	442
<i>Code-Switching for Enhancing NMT with Pre-Specified Translation</i> Kai Song, Yue Zhang, Heng Yu, Weihua Luo, Kun Wang and Min Zhang	449
<i>Aligning Vector-spaces with Noisy Supervised Lexicon</i> Noa Yehezkel Lubin, Jacob Goldberger and Yoav Goldberg	460
<i>Understanding and Improving Hidden Representations for Neural Machine Translation</i> Guanlin Li, Lemao Liu, Xintong Li, Conghui Zhu, Tiejun Zhao and Shuming Shi	466
<i>Content Differences in Syntactic and Semantic Representation</i> Daniel Hershcovich, Omri Abend and Ari Rappoport	478
<i>Attentive Mimicking: Better Word Embeddings by Attending to Informative Contexts</i> Timo Schick and Hinrich Schütze	489
<i>Evaluating Style Transfer for Text</i> Remi Mir, Bjarke Felbo, Nick Obradovich and Iyad Rahwan	495
<i>Big BiRD: A Large, Fine-Grained, Bigram Relatedness Dataset for Examining Semantic Composition</i> Shima Asaadi, Saif Mohammad and Svetlana Kiritchenko	505
<i>Outlier Detection for Improved Data Quality and Diversity in Dialog Systems</i> Stefan Larson, Anish Mahendran, Andrew Lee, Jonathan K. Kummerfeld, Parker Hill, Michael A. Laurenzano, Johann Hauswald, Lingjia Tang and Jason Mars	517

<i>Asking the Right Question: Inferring Advice-Seeking Intentions from Personal Narratives</i> Liye Fu, Jonathan P. Chang and Cristian Danescu-Niculescu-Mizil	528
<i>Seeing Things from a Different Angle: Discovering Diverse Perspectives about Claims</i> Sihao Chen, Daniel Khashabi, Wenpeng Yin, Chris Callison-Burch and Dan Roth	542
<i>IMHO Fine-Tuning Improves Claim Detection</i> Tuhin Chakrabarty, Christopher Hidey and Kathy McKeown	558
<i>Joint Multiple Intent Detection and Slot Labeling for Goal-Oriented Dialog</i> Rashmi Gangadharaiiah and Balakrishnan Narayanaswamy	564
<i>CITE: A Corpus of Image-Text Discourse Relations</i> Malihe Alikhani, Sreyasi Nag Chowdhury, Gerard de Melo and Matthew Stone	570
<i>Improving Dialogue State Tracking by Discerning the Relevant Context</i> Sanuj Sharma, Prafulla Kumar Choubey and Ruihong Huang	576
<i>CLEVR-Dialog: A Diagnostic Dataset for Multi-Round Reasoning in Visual Dialog</i> Satwik Kottur, José M. F. Moura, Devi Parikh, Dhruv Batra and Marcus Rohrbach	582
<i>Learning Outside the Box: Discourse-level Features Improve Metaphor Identification</i> Jesse Mu, Helen Yannakoudakis and Ekaterina Shutova	596
<i>Detection of Abusive Language: the Problem of Biased Datasets</i> Michael Wiegand, Josef Ruppenhofer and Thomas Kleinbauer	602
<i>Lipstick on a Pig: Debiasing Methods Cover up Systematic Gender Biases in Word Embeddings But do not Remove Them</i> Hila Gonen and Yoav Goldberg	609
<i>Black is to Criminal as Caucasian is to Police: Detecting and Removing Multiclass Bias in Word Embeddings</i> Thomas Manzini, Lim Yao Chong, Alan W. Black and Yulia Tsvetkov	615
<i>On Measuring Social Biases in Sentence Encoders</i> Chandler May, Alex Wang, Shikha Bordia, Samuel R. Bowman and Rachel Rudinger	622
<i>Gender Bias in Contextualized Word Embeddings</i> Jieyu Zhao, Tianlu Wang, Mark Yatskar, Ryan Cotterell, Vicente Ordonez and Kai-Wei Chang	629
<i>Combining Sentiment Lexica with a Multi-View Variational Autoencoder</i> Alexander Miserlis Hoyle, Lawrence Wolf-Sonkin, Hanna Wallach, Ryan Cotterell and Isabelle Augenstein	635
<i>Enhancing Opinion Role Labeling with Semantic-Aware Word Representations from Semantic Role Labeling</i> Meishan Zhang, Peili Liang and Guohong Fu	641
<i>Frowning Frodo, Wincing Leia, and a Seriously Great Friendship: Learning to Classify Emotional Relationships of Fictional Characters</i> Evgeny Kim and Roman Klinger	647
<i>Generalizing Unmasking for Short Texts</i> Janek Bevendorff, Benno Stein, Matthias Hagen and Martin Potthast	654

<i>Adversarial Training for Satire Detection: Controlling for Confounding Variables</i> Robert McHardy, Heike Adel and Roman Klinger	660
<i>Keyphrase Generation: A Text Summarization Struggle</i> Erion Çano and Ondřej Bojar	666
<i>SEQ³: Differentiable Sequence-to-Sequence-to-Sequence Autoencoder for Unsupervised Abstractive Sentence Compression</i> Christos Baziotis, Ion Androutsopoulos, Ioannis Konstas and Alexandros Potamianos	673
<i>Crowdsourcing Lightweight Pyramids for Manual Summary Evaluation</i> Ori Shapira, David Gabay, Yang Gao, Hadar Ronen, Ramakanth Pasunuru, Mohit Bansal, Yael Amsterdamer and Ido Dagan	682
<i>Serial Recall Effects in Neural Language Modeling</i> Hassan Hajjipoor, Hadi Amiri, Maseud Rahgozar and Farhad Oroumchian	688
<i>Fast Concept Mention Grouping for Concept Map-based Multi-Document Summarization</i> Tobias Falke and Iryna Gurevych	695
<i>Syntax-aware Neural Semantic Role Labeling with Supertags</i> Jungo Kasai, Dan Friedman, Robert Frank, Dragomir Radev and Owen Rambow	701
<i>Left-to-Right Dependency Parsing with Pointer Networks</i> Daniel Fernández-González and Carlos Gómez-Rodríguez	710
<i>Viable Dependency Parsing as Sequence Labeling</i> Michalina Strzyz, David Vilares and Carlos Gómez-Rodríguez	717
<i>Pooled Contextualized Embeddings for Named Entity Recognition</i> Alan Akbik, Tanja Bergmann and Roland Vollgraf	724
<i>Better Modeling of Incomplete Annotations for Named Entity Recognition</i> Zhanming Jie, Pengjun Xie, Wei Lu, Ruixue Ding and Linlin Li	729
<i>Event Detection without Triggers</i> Shulin Liu, Yang Li, Feng Zhang, Tao Yang and Xinpeng Zhou	735
<i>Sub-event detection from twitter streams as a sequence labeling problem</i> Giannis Bekoulis, Johannes Deleu, Thomas Demeester and Chris Develder	745
<i>GraphIE: A Graph-Based Framework for Information Extraction</i> Yujie Qian, Enrico Santus, Zhijing Jin, Jiang Guo and Regina Barzilay	751
<i>OpenKI: Integrating Open Information Extraction and Knowledge Bases with Relation Inference</i> Dongxu Zhang, Subhabrata Mukherjee, Colin Lockard, Luna Dong and Andrew McCallum ...	762
<i>Imposing Label-Relational Inductive Bias for Extremely Fine-Grained Entity Typing</i> Wenhan Xiong, Jiawei Wu, Deren Lei, Mo Yu, Shiyu Chang, Xiaoxiao Guo and William Yang Wang	773
<i>Improving Event Coreference Resolution by Learning Argument Compatibility from Unlabeled Data</i> Yin Jou Huang, Jing Lu, Sadao Kurohashi and Vincent Ng	785
<i>Sentence Embedding Alignment for Lifelong Relation Extraction</i> Hong Wang, Wenhan Xiong, Mo Yu, Xiaoxiao Guo, Shiyu Chang and William Yang Wang ...	796

<i>Description-Based Zero-shot Fine-Grained Entity Typing</i> Rasha Obeidat, Xiaoli Fern, Hamed Shahbazi and Prasad Tadepalli	807
<i>Adversarial Decomposition of Text Representation</i> Alexey Romanov, Anna Rumshisky, Anna Rogers and David Donahue	815
<i>PoMo: Generating Entity-Specific Post-Modifiers in Context</i> Jun Seok Kang, Robert Logan, Zewei Chu, Yang Chen, Dheeru Dua, Kevin Gimpel, Sameer Singh and Niranjana Balasubramanian	826
<i>Improved Lexically Constrained Decoding for Translation and Monolingual Rewriting</i> J. Edward Hu, Huda Khayrallah, Ryan Culkin, Patrick Xia, Tongfei Chen, Matt Post and Benjamin Van Durme	839
<i>Courteously Yours: Inducing courteous behavior in Customer Care responses using Reinforced Pointer Generator Network</i> Hitesh Golchha, Mauajama Firdaus, Asif Ekbal and Pushpak Bhattacharyya	851
<i>How to Avoid Sentences Spelling Boring? Towards a Neural Approach to Unsupervised Metaphor Generation</i> Zhiwei Yu and Xiaojun Wan	861
<i>Incorporating Context and External Knowledge for Pronoun Coreference Resolution</i> Hongming Zhang, Yan Song and Yangqiu Song	872
<i>Unsupervised Deep Structured Semantic Models for Commonsense Reasoning</i> Shuohang Wang, Sheng Zhang, Yelong Shen, Xiaodong Liu, Jingjing Liu, Jianfeng Gao and Jing Jiang	882
<i>Recovering dropped pronouns in Chinese conversations via modeling their referents</i> Jingxuan Yang, Jianzhuo Tong, Si Li, Sheng Gao, Jun Guo and Nianwen Xue	892
<i>The problem with probabilistic DAG automata for semantic graphs</i> Ieva Vasiljeva, Sorcha Gilroy and Adam Lopez	902
<i>A Systematic Study of Leveraging Subword Information for Learning Word Representations</i> Yi Zhu, Ivan Vulić and Anna Korhonen	912
<i>Better Word Embeddings by Disentangling Contextual n-Gram Information</i> Prakhar Gupta, Matteo Pagliardini and Martin Jaggi	933
<i>Integration of Knowledge Graph Embedding Into Topic Modeling with Hierarchical Dirichlet Process</i> Dingcheng Li, Siamak Zamani, Jingyuan Zhang and Ping Li	940
<i>Correlation Coefficients and Semantic Textual Similarity</i> Vitalii Zhelezniak, Aleksandar Savkov, April Shen and Nils Hammerla	951
<i>Generating Token-Level Explanations for Natural Language Inference</i> James Thorne, Andreas Vlachos, Christos Christodoulopoulos and Arpit Mittal	963
<i>Strong Baselines for Complex Word Identification across Multiple Languages</i> Pierre Finamore, Elisabeth Fritzsche, Daniel King, Alison Snelyd, Aneeq Ur Rehman, Fernando Alva-Manchego and Andreas Vlachos	970
<i>Adaptive Convolution for Multi-Relational Learning</i> Xiaotian Jiang, Quan Wang and Bin Wang	978

<i>Graph Pattern Entity Ranking Model for Knowledge Graph Completion</i> Takuma Ebisu and Ryutaro Ichise	988
<i>Adversarial Training for Weakly Supervised Event Detection</i> Xiaozhi Wang, Xu Han, Zhiyuan Liu, Maosong Sun and Peng Li	998
<i>A Submodular Feature-Aware Framework for Label Subset Selection in Extreme Classification Problems</i> Elham J. Barezi, Ian D. Wood, Pascale Fung and Hamid R. Rabiee	1009
<i>Relation Extraction with Temporal Reasoning Based on Memory Augmented Distant Supervision</i> Jianhao Yan, Lin He, Ruqin Huang, Jian Li and Ying Liu	1019
<i>Integrating Semantic Knowledge to Tackle Zero-shot Text Classification</i> Jingqing Zhang, Piyawat Lertvittayakumjorn and Yike Guo	1031
<i>Word-Node2Vec: Improving Word Embedding with Document-Level Non-Local Word Co-occurrences</i> Procheta Sen, Debasis Ganguly and Gareth Jones	1041
<i>Cross-Topic Distributional Semantic Representations Via Unsupervised Mappings</i> Eleftheria Briakou, Nikos Athanasiou and Alexandros Potamianos	1052
<i>What just happened? Evaluating retrofitted distributional word vectors</i> Dmetri Hayes	1062
<i>Linguistic Knowledge and Transferability of Contextual Representations</i> Nelson F. Liu, Matt Gardner, Yonatan Belinkov, Matthew E. Peters and Noah A. Smith	1073
<i>Mutual Information Maximization for Simple and Accurate Part-Of-Speech Induction</i> Karl Stratos	1095
<i>Unsupervised Recurrent Neural Network Grammars</i> Yoon Kim, Alexander Rush, Lei Yu, Adhiguna Kuncoro, Chris Dyer and Gábor Melis	1105
<i>Cooperative Learning of Disjoint Syntax and Semantics</i> Serhii Havrylov, Germán Kruszewski and Armand Joulin	1118
<i>Unsupervised Latent Tree Induction with Deep Inside-Outside Recursive Auto-Encoders</i> Andrew Drozdov, Patrick Verga, Mohit Yadav, Mohit Iyyer and Andrew McCallum	1129
<i>Knowledge-Augmented Language Model and Its Application to Unsupervised Named-Entity Recognition</i> Angli Liu, Jingfei Du and Veselin Stoyanov	1142
<i>Syntax-Enhanced Neural Machine Translation with Syntax-Aware Word Representations</i> Meishan Zhang, Zhenghua Li, Guohong Fu and Min Zhang	1151
<i>Competence-based Curriculum Learning for Neural Machine Translation</i> Emmanouil Antonios Platanios, Otilia Stretcu, Graham Neubig, Barnabas Poczos and Tom Mitchell	1162
<i>Extract and Edit: An Alternative to Back-Translation for Unsupervised Neural Machine Translation</i> Jiawei Wu, Xin Wang and William Yang Wang	1173
<i>Consistency by Agreement in Zero-Shot Neural Machine Translation</i> Maruan Al-Shedivat and Ankur Parikh	1184

<i>Modeling Recurrence for Transformer</i>	
Jie Hao, Xing Wang, Baosong Yang, Longyue Wang, Jinfeng Zhang and Zhaopeng Tu	1198
<i>Rethinking Action Spaces for Reinforcement Learning in End-to-end Dialog Agents with Latent Variable Models</i>	
Tiancheng Zhao, Kaige Xie and Maxine Eskenazi	1208
<i>Skeleton-to-Response: Dialogue Generation Guided by Retrieval Memory</i>	
Deng Cai, Yan Wang, Wei Bi, Zhaopeng Tu, Xiaojiang Liu, Wai Lam and Shuming Shi	1219
<i>Jointly Optimizing Diversity and Relevance in Neural Response Generation</i>	
Xiang Gao, Sungjin Lee, Yizhe Zhang, Chris Brockett, Michel Galley, Jianfeng Gao and Bill Dolan	1229
<i>Disentangling Language and Knowledge in Task-Oriented Dialogs</i>	
Dinesh Raghu, Nikhil Gupta and	1239
<i>Tensorized Self-Attention: Efficiently Modeling Pairwise and Global Dependencies Together</i>	
Tao Shen, Tianyi Zhou, Guodong Long, Jing Jiang and Chengqi Zhang	1256
<i>WiC: the Word-in-Context Dataset for Evaluating Context-Sensitive Meaning Representations</i>	
Mohammad Taher Pilehvar and Jose Camacho-Collados	1267
<i>Does My Rebuttal Matter? Insights from a Major NLP Conference</i>	
Yang Gao, Steffen Eger, Ilia Kuznetsov, Iryna Gurevych and Yusuke Miyao	1274
<i>CASTING LIGHT ON INVISIBLE CITIES: COMPUTATIONALLY ENGAGING WITH LITERARY CRITICISM</i>	
Shufan Wang and Mohit Iyyer	1291
<i>PAWS: Paraphrase Adversaries from Word Scrambling</i>	
Yuan Zhang, Jason Baldridge and Luheng He	1298
<i>Cross-Corpora Evaluation and Analysis of Grammatical Error Correction Models — Is Single-Corpus Evaluation Enough?</i>	
Masato Mita, Tomoya Mizumoto, Masahiro Kaneko, Ryo Nagata and Kentaro Inui	1309
<i>Star-Transformer</i>	
Qipeng Guo, Xipeng Qiu, Pengfei Liu, Yunfan Shao, Xiangyang Xue and Zheng Zhang	1315
<i>Adaptation of Hierarchical Structured Models for Speech Act Recognition in Asynchronous Conversation</i>	
Tasnim Mohiuddin, Thanh-Tung Nguyen and Shafiq Joty	1326
<i>From legal to technical concept: Towards an automated classification of German political Twitter postings as criminal offenses</i>	
Frederike Zufall, Tobias Horsmann and Torsten Zesch	1337
<i>Joint Multi-Label Attention Networks for Social Text Annotation</i>	
Hang Dong, Wei Wang, Kaizhu Huang and Frans Coenen	1348
<i>Multi-Channel Convolutional Neural Network for Twitter Emotion and Sentiment Recognition</i>	
Jumayel Islam, Robert E. Mercer and Lu Xiao	1355
<i>Detecting Cybersecurity Events from Noisy Short Text</i>	
Semih Yagcioglu, Mehmet saygin Seyfioglu, Begum Citamak, Batuhan Bardak, Seren Guldamlasioglu, Azmi Yuksel and Emin Islam Tatli	1366

<i>White-to-Black: Efficient Distillation of Black-Box Adversarial Attacks</i> Yotam Gil, Yoav Chai, Or Gorodissky and Jonathan Berant	1373
<i>Analyzing the Perceived Severity of Cybersecurity Threats Reported on Social Media</i> Shi Zong, Alan Ritter, Graham Mueller and Evan Wright	1380
<i>Fake News Detection using Deep Markov Random Fields</i> Duc Minh Nguyen, Tien Huu Do, Robert Calderbank and Nikos Deligiannis	1391
<i>Issue Framing in Online Discussion Fora</i> Mareike Hartmann, Tallulah Jansen, Isabelle Augenstein and Anders Søgaard	1401
<i>Vector of Locally Aggregated Embeddings for Text Representation</i> Hadi Amiri and Mitra Mohtarami	1408
<i>Predicting the Type and Target of Offensive Posts in Social Media</i> Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra and Ritesh Kumar 1415	
<i>Biomedical Event Extraction based on Knowledge-driven Tree-LSTM</i> Diya Li, Lifu Huang, Heng Ji and Jiawei Han	1421
<i>Detecting cognitive impairments by agreeing on interpretations of linguistic features</i> Zining Zhu, Jekaterina Novikova and Frank Rudzicz	1431
<i>Relation Extraction using Explicit Context Conditioning</i> Gaurav Singh and Parminder Bhatia	1442
<i>Conversation Model Fine-Tuning for Classifying Client Utterances in Counseling Dialogues</i> Sungjoon Park, Donghyun Kim and Alice Oh	1448
<i>Using Similarity Measures to Select Pretraining Data for NER</i> Xiang Dai, Sarvnaz Karimi, Ben Hachey and Cecile Paris	1460
<i>Predicting Annotation Difficulty to Improve Task Routing and Model Performance for Biomedical Information Extraction</i> Yinfei Yang, Oshin Agarwal, Chris Tar, Byron C. Wallace and Ani Nenkova	1471
<i>Detecting Depression in Social Media using Fine-Grained Emotions</i> Mario Ezra Aragon, Adrian Pastor Lopez Monroy, Luis Carlos Gonzalez Gurrola and Manuel Montes-y-Gomez	1481
<i>A Silver Standard Corpus of Human Phenotype-Gene Relations</i> Diana Sousa, Andre Lamurias and Francisco M Couto	1487
<i>Improving Lemmatization of Non-Standard Languages with Joint Learning</i> Enrique Manjavacas, Ákos Kádár and Mike Kestemont	1493
<i>One Size Does Not Fit All: Comparing NMT Representations of Different Granularities</i> Nadir Durrani, Fahim Dalvi, Hassan Sajjad, Yonatan Belinkov and Preslav Nakov	1504
<i>A Simple Joint Model for Improved Contextual Neural Lemmatization</i> Chaitanya Malaviya, Shijie Wu and Ryan Cotterell	1517
<i>A Probabilistic Generative Model of Linguistic Typology</i> Johannes Bjerva, Yova Kementchedjieva, Ryan Cotterell and Isabelle Augenstein	1529

<i>Quantifying the morphosyntactic content of Brown Clusters</i> Manuel Ciosici, Leon Derczynski and Ira Assent	1541
<i>Analyzing Bayesian Crosslingual Transfer in Topic Models</i> Shudong Hao and Michael J. Paul	1551
<i>Recursive Subtree Composition in LSTM-Based Dependency Parsing</i> Miryam de Lhoneux, Miguel Ballesteros and Joakim Nivre	1566
<i>Cross-lingual CCG Induction</i> Kilian Evang	1577
<i>Density Matching for Bilingual Word Embedding</i> Chunting Zhou, Xuezhe Ma, Di Wang and Graham Neubig	1588
<i>Cross-Lingual Alignment of Contextual Word Embeddings, with Applications to Zero-shot Dependency Parsing</i> Tal Schuster, Ori Ram, Regina Barzilay and Amir Globerson	1599
<i>Early Rumour Detection</i> Kaimin Zhou, Chang Shu, Binyang Li and Jey Han Lau	1614
<i>Microblog Hashtag Generation via Encoding Conversation Contexts</i> Yue Wang, Jing Li, Irwin King, Michael R. Lyu and Shuming Shi	1624
<i>Text Processing Like Humans Do: Visually Attacking and Shielding NLP Systems</i> Steffen Eger, Gözde Gül Şahin, Andreas Rücklé, Ji-Ung Lee, Claudia Schulz, Mohsen Mesgar, Krishnkant Swarnkar, Edwin Simpson and Iryna Gurevych	1634
<i>Something’s Brewing! Early Prediction of Controversy-causing Posts from Discussion Features</i> Jack Hessel and Lillian Lee	1648
<i>No Permanent Friends or Enemies: Tracking Relationships between Nations from News</i> Xiaochuang Han, Eunsol Choi and Chenhao Tan	1660
<i>Improving Human Text Comprehension through Semi-Markov CRF-based Neural Section Title Generation</i> Sebastian Gehrmann, Steven Layne and Franck Dernoncourt	1677
<i>Unifying Human and Statistical Evaluation for Natural Language Generation</i> Tatsunori Hashimoto, Hugh Zhang and Percy Liang	1689
<i>What makes a good conversation? How controllable attributes affect human judgments</i> Abigail See, Stephen Roller, Douwe Kiela and Jason Weston	1702
<i>An Empirical Investigation of Global and Local Normalization for Recurrent Neural Sequence Models Using a Continuous Relaxation to Beam Search</i> Kartik Goyal, Chris Dyer and Taylor Berg-Kirkpatrick	1724
<i>Pun Generation with Surprise</i> He He, Nanyun Peng and Percy Liang	1734
<i>Single Document Summarization as Tree Induction</i> Yang Liu, Ivan Titov and Mirella Lapata	1745

<i>Fixed That for You: Generating Contrastive Claims with Semantic Edits</i> Christopher Hidey and Kathy McKeown	1756
<i>Box of Lies: Multimodal Deception Detection in Dialogues</i> Felix Soldner, Verónica Pérez-Rosas and Rada Mihalcea	1768
<i>A Crowdsourced Corpus of Multiple Judgments and Disagreement on Anaphoric Interpretation</i> Massimo Poesio, Jon Chamberlain, Silviu Paun, Juntao Yu, Alexandra Uma and Udo Kruschwitz 1778	
<i>A Streamlined Method for Sourcing Discourse-level Argumentation Annotations from the Crowd</i> Tristan Miller, Maria Sukhareva and Iryna Gurevych	1790
<i>Unsupervised Dialog Structure Learning</i> Weiyang Shi, Tiancheng Zhao and Zhou Yu	1797
<i>Modeling Document-level Causal Structures for Event Causal Relation Identification</i> Lei Gao, Prafulla Kumar Choubey and Ruihong Huang	1808
<i>Hierarchical User and Item Representation with Three-Tier Attention for Recommendation</i> Chuhan Wu, Fangzhao Wu, Junxin Liu and Yongfeng Huang	1818
<i>Text Similarity Estimation Based on Word Embeddings and Matrix Norms for Targeted Marketing</i> Tim vor der Brück and Marc Pouly	1827
<i>Glocal: Incorporating Global Information in Local Convolution for Keyphrase Extraction</i> Animesh Prasad and Min-Yen Kan	1837
<i>A Study of Latent Structured Prediction Approaches to Passage Reranking</i> Iryna Haponchyk and Alessandro Moschitti	1847
<i>Combining Distant and Direct Supervision for Neural Relation Extraction</i> Iz Beltagy, Kyle Lo and Waleed Ammar	1858
<i>Tweet Stance Detection Using an Attention based Neural Ensemble Model</i> Umme Aymun Siddiqua, Abu Nowshed Chy and Masaki Aono	1868
<i>Word Embedding-Based Automatic MT Evaluation Metric using Word Position Information</i> Hiroshi Echizen'ya, Kenji Araki and Eduard Hovy	1874
<i>Learning to Stop in Structured Prediction for Neural Machine Translation</i> Mingbo Ma, Renjie Zheng and Liang Huang	1884
<i>Learning Unsupervised Multilingual Word Embeddings with Incremental Multilingual Hubs</i> Geert Heyman, Bregt Verreet, Ivan Vulić and Marie-Francine Moens	1890
<i>Curriculum Learning for Domain Adaptation in Neural Machine Translation</i> Xuan Zhang, Pamela Shapiro, Gaurav Kumar, Paul McNamee, Marine Carpuat and Kevin Duh	1903
<i>Improving Robustness of Machine Translation with Synthetic Noise</i> Vaibhav Vaibhav, Sumeet Singh, Craig Stewart and Graham Neubig	1916
<i>Non-Parametric Adaptation for Neural Machine Translation</i> Ankur Bapna and Orhan Firat	1921

<i>Online Distilling from Checkpoints for Neural Machine Translation</i> Hao-Ran Wei, Shujian Huang, Ran Wang, Xin-Yu Dai and Jiajun Chen	1932
<i>Value-based Search in Execution Space for Mapping Instructions to Programs</i> Dor Muhlgay, Jonathan Herzig and Jonathan Berant	1942
<i>VQD: Visual Query Detection In Natural Scenes</i> Manoj Acharya, Karan Jariwala and Christopher Kanan	1955
<i>Improving Natural Language Interaction with Robots Using Advice</i> Nikhil Mehta and Dan Goldwasser	1962
<i>Generating Knowledge Graph Paths from Textual Definitions using Sequence-to-Sequence Models</i> Victor Prokhorov, Mohammad Taher Pilehvar and Nigel Collier	1968
<i>Shifting the Baseline: Single Modality Performance on Visual Navigation & QA</i> Jesse Thomason, Daniel Gordon and Yonatan Bisk	1977
<i>ExCL: Extractive Clip Localization Using Natural Language Descriptions</i> Soham Ghosh, Anuva Agarwal, Zarana Parekh and Alexander Hauptmann	1984
<i>Detecting dementia in Mandarin Chinese using transfer learning from a parallel corpus</i> Bai Li, Yi-Te Hsu and Frank Rudzicz	1991
<i>Cross-lingual Visual Verb Sense Disambiguation</i> Spandana Gella, Desmond Elliott and Frank Keller	1998
<i>Subword-Level Language Identification for Intra-Word Code-Switching</i> Manuel Mager, Özlem Çetinoğlu and Katharina Kann	2005
<i>MuST-C: a Multilingual Speech Translation Corpus</i> Mattia A. Di Gangi, Roldano Cattoni, Luisa Bentivogli, Matteo Negri and Marco Turchi	2012
<i>Contextualization of Morphological Inflection</i> Ekaterina Vylomova, Ryan Cotterell, Trevor Cohn, Timothy Baldwin and Jason Eisner	2018
<i>A Robust Abstractive System for Cross-Lingual Summarization</i> Jessica Ouyang, Boya Song and Kathy McKeown	2025
<i>Improving Neural Machine Translation with Neural Syntactic Distance</i> Chunpeng Ma, Akihiro Tamura, Masao Utiyama, Eiichiro Sumita and Tiejun Zhao	2032
<i>Measuring Immediate Adaptation Performance for Neural Machine Translation</i> Patrick Simianer, Joern Wuebker and John DeNero	2038
<i>Differentiable Sampling with Flexible Reference Word Order for Neural Machine Translation</i> Weijia Xu, Xing Niu and Marine Carpuat	2047
<i>Reinforcement Learning based Curriculum Optimization for Neural Machine Translation</i> Gaurav Kumar, George Foster, Colin Cherry and Maxim Krikun	2054
<i>Overcoming Catastrophic Forgetting During Domain Adaptation of Neural Machine Translation</i> Brian Thompson, Jeremy Gwinnup, Huda Khayrallah, Kevin Duh and Philipp Koehn	2062
<i>Short-Term Meaning Shift: A Distributional Exploration</i> Marco Del Tredici, Raquel Fernández and Gemma Boleda	2069

<i>Detecting Derogatory Compounds – An Unsupervised Approach</i> Michael Wiegand, Maximilian Wolf and Josef Ruppenhofer	2076
<i>Personalized Neural Embeddings for Collaborative Filtering with Text</i> Guangneng Hu	2082
<i>An Embarrassingly Simple Approach for Transfer Learning from Pretrained Language Models</i> Alexandra Chronopoulou, Christos Baziotis and Alexandros Potamianos	2089
<i>Incorporating Emoji Descriptions Improves Tweet Classification</i> Abhishek Singh, Eduardo Blanco and Wei Jin	2096
<i>Modeling Personal Biases in Language Use by Inducing Personalized Word Embeddings</i> Daisuke Oba, Naoki Yoshinaga, Shoetsu Sato, Satoshi Akasaki and Masashi Toyoda	2102
<i>Multi-Task Ordinal Regression for Jointly Predicting the Trustworthiness and the Leading Political Ideology of News Media</i> Ramy Baly, Georgi Karadzhov, Abdelrhman Saleh, James Glass and Preslav Nakov	2109
<i>Joint Detection and Location of English Puns</i> Yanyan Zou and Wei Lu	2117
<i>Harry Potter and the Action Prediction Challenge from Natural Language</i> David Vilares and Carlos Gómez-Rodríguez	2124
<i>Argument Mining for Understanding Peer Reviews</i> Xinyu Hua, Mitko Nikolov, Nikhil Badugu and Lu Wang	2131
<i>An annotated dataset of literary entities</i> David Bamman, Sejal Popat and Sheng Shen	2138
<i>Abusive Language Detection with Graph Convolutional Networks</i> Pushkar Mishra, Marco Del Tredici, Helen Yannakoudakis and Ekaterina Shutova	2145
<i>On the Importance of Distinguishing Word Meaning Representations: A Case Study on Reverse Dictionary Mapping</i> Mohammad Taher Pilehvar	2151
<i>Factorising AMR generation through syntax</i> Kris Cao and Stephen Clark	2157
<i>A Crowdsourced Frame Disambiguation Corpus with Ambiguity</i> Anca Dumitrache, Lora Aroyo and Chris Welty	2164
<i>Inoculation by Fine-Tuning: A Method for Analyzing Challenge Datasets</i> Nelson F. Liu, Roy Schwartz and Noah A. Smith	2171
<i>A Capsule Network-based Embedding Model for Knowledge Graph Completion and Search Personalization</i> Dai Quoc Nguyen, Thanh Vu, Tu Dinh Nguyen, Dat Quoc Nguyen and Dinh Phung	2180
<i>Partial Or Complete, That's The Question</i> Qiang Ning, Hangfeng He, Chuchu Fan and Dan Roth	2190
<i>Sequential Attention with Keyword Mask Model for Community-based Question Answering</i> Jianxin Yang, Wenge Rong, Libin Shi and Zhang Xiong	2201

<i>Simple Attention-Based Representation Learning for Ranking Short Social Media Posts</i> Peng Shi, Jinfeng Rao and Jimmy Lin	2212
<i>AttentiveChecker: A Bi-Directional Attention Flow Mechanism for Fact Verification</i> Santosh Tokala, Vishal G, Avirup Saha and Niloy Ganguly	2218
<i>Practical, Efficient, and Customizable Active Learning for Named Entity Recognition in the Digital Humanities</i> Alexander Erdmann, David Joseph Wrisley, Benjamin Allen, Christopher Brown, Sophie Cohen-Bodénès, Micha Elsner, Yukun Feng, Brian Joseph, Béatrice Joyeux-Prunel and Marie-Catherine de Marneffe	2223
<i>Doc2hash: Learning Discrete Latent variables for Documents Retrieval</i> Yifei Zhang and Hao Zhu	2235
<i>Evaluating Text GANs as Language Models</i> Guy Tevet, Gavriel Habib, Vered Shwartz and Jonathan Berant	2241
<i>Latent Code and Text-based Generative Adversarial Networks for Soft-text Generation</i> Md Akmal Haidar, Mehdi Rezagholizadeh, Alan Do Omri and Ahmad Rashid	2248
<i>Neural Text Generation from Rich Semantic Representations</i> Valerie Hajdik, Jan Buys, Michael Wayne Goodman and Emily M. Bender	2259
<i>Step-by-Step: Separating Planning from Realization in Neural Data-to-Text Generation</i> Amit Moryossef, Yoav Goldberg and Ido Dagan	2267
<i>Evaluating Rewards for Question Generation Models</i> Tom Hosking and Sebastian Riedel	2278
<i>Text Generation from Knowledge Graphs with Graph Transformers</i> Rik Koncel-Kedziorski, Dhanush Bekal, Yi Luan, Mirella Lapata and Hannaneh Hajishirzi ..	2284
<i>Open Information Extraction from Question-Answer Pairs</i> Nikita Bhutani, Yoshihiko Suhara, Wang-Chiew Tan, Alon Halevy and H. V. Jagadish	2294
<i>Question Answering by Reasoning Across Documents with Graph Convolutional Networks</i> Nicola De Cao, Wilker Aziz and Ivan Titov	2306
<i>A Qualitative Comparison of CoQA, SQuAD 2.0 and QuAC</i> Mark Yatskar	2318
<i>BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis</i> Hu Xu, Bing Liu, Lei Shu and Philip Yu	2324
<i>Old is Gold: Linguistic Driven Approach for Entity and Relation Linking of Short Text</i> Ahmad Sakor, Isaiah Onando Mulang [†] , Kuldeep Singh, Saeedeh Shekarpour, Maria Esther Vidal, Jens Lehmann and Sören Auer	2336
<i>Be Consistent! Improving Procedural Text Comprehension using Label Consistency</i> Xinya Du, Bhavana Dalvi, Niket Tandon, Antoine Bosselut, Wen-tau Yih, Peter Clark and Claire Cardie	2347
<i>MathQA: Towards Interpretable Math Word Problem Solving with Operation-Based Formalisms</i> Aida Amini, Saadia Gabriel, Shanchuan Lin, Rik Koncel-Kedziorski, Yejin Choi and Hannaneh Hajishirzi	2357

<i>DROP: A Reading Comprehension Benchmark Requiring Discrete Reasoning Over Paragraphs</i> Dheeru Dua, Yizhong Wang, Pradeep Dasigi, Gabriel Stanovsky, Sameer Singh and Matt Gardner	2368
<i>An Encoding Strategy Based Word-Character LSTM for Chinese NER</i> Wei Liu, Tongge Xu, Qinghua Xu, Jiayu Song and Yueran Zu	2379
<i>Highly Effective Arabic Diacritization using Sequence to Sequence Modeling</i> Hamdy Mubarak, Ahmed Abdelali, Hassan Sajjad, Younes Samih and Kareem Darwish	2390
<i>SC-LSTM: Learning Task-Specific Representations in Multi-Task Learning for Sequence Labeling</i> Peng Lu, Ting Bai and Philippe Langlais	2396
<i>Learning to Denoise Distantly-Labeled Data for Entity Typing</i> Yasumasa Onoe and Greg Durrett	2407
<i>A Simple and Robust Approach to Detecting Subject-Verb Agreement Errors</i> Simon Flachs, Ophélie Lacroix, Marek Rei, Helen Yannakoudakis and Anders Søgaard	2418
<i>A Grounded Unsupervised Universal Part-of-Speech Tagger for Low-Resource Languages</i> Ronald Cardenas, Ying Lin, Heng Ji and Jonathan May	2428
<i>On Difficulties of Cross-Lingual Transfer with Order Differences: A Case Study on Dependency Parsing</i> Wasi Ahmad, Zhisong Zhang, Xuezhe Ma, Eduard Hovy, Kai-Wei Chang and Nanyun Peng	2440
<i>A Multi-Task Approach for Disentangling Syntax and Semantics in Sentence Representations</i> Mingda Chen, Qingming Tang, Sam Wiseman and Kevin Gimpel	2453
<i>Self-Discriminative Learning for Unsupervised Document Embedding</i> Hong-You Chen, Chin-Hua Hu, Leila Wehbe and Shou-de Lin	2465
<i>Adaptive Convolution for Text Classification</i> Byung-Ju Choi, Jun-Hyung Park and SangKeun Lee	2475
<i>Zero-Shot Cross-Lingual Opinion Target Extraction</i> Soufian Jebbara and Philipp Cimiano	2486
<i>Adversarial Category Alignment Network for Cross-domain Sentiment Classification</i> Xiaoye Qu, Zhikang Zou, Yu Cheng, Yang Yang and Pan Zhou	2496
<i>Target-oriented Opinion Words Extraction with Target-fused Neural Sequence Labeling</i> Zhifang Fan, Zhen Wu, Xin-Yu Dai, Shujian Huang and Jiajun Chen	2509
<i>Abstractive Summarization of Reddit Posts with Multi-level Memory Networks</i> Byeongchang Kim, Hyunwoo Kim and Gunhee Kim	2519
<i>Automatic learner summary assessment for reading comprehension</i> Menglin Xia, Ekaterina Kochmar and Ted Briscoe	2532
<i>Data-efficient Neural Text Compression with Interactive Learning</i> Avinesh P.V.S and Christian M. Meyer	2543
<i>Text Generation with Exemplar-based Adaptive Decoding</i> Hao Peng, Ankur Parikh, Manaal Faruqui, Bhuwan Dhingra and Dipanjan Das	2555

<i>Guiding Extractive Summarization with Question-Answering Rewards</i> Kristjan Arumae and Fei Liu	2566
<i>Beyond task success: A closer look at jointly learning to see, ask, and GuessWhat</i> Ravi Shekhar, Aashish Venkatesh, Tim Baumgärtner, Elia Bruni, Barbara Plank, Raffaella Bernardi and Raquel Fernández	2578
<i>The World in My Mind: Visual Dialog with Adversarial Multi-modal Feature Encoding</i> Yiqun Yao, Jiaming Xu and Bo Xu	2588
<i>Strong and Simple Baselines for Multimodal Utterance Embeddings</i> Paul Pu Liang, Yao Chong Lim, Yao-Hung Hubert Tsai, Ruslan Salakhutdinov and Louis-Philippe Morency	2599
<i>Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout</i> Hao Tan, Licheng Yu and Mohit Bansal	2610
<i>Towards Content Transfer through Grounded Text Generation</i> Shrimai Prabhumoye, Chris Quirk and Michel Galley	2622
<i>Improving Machine Reading Comprehension with General Reading Strategies</i> Kai Sun, Dian Yu, Dong Yu and Claire Cardie	2633
<i>Multi-task Learning with Sample Re-weighting for Machine Reading Comprehension</i> Yichong Xu, Xiaodong Liu, Yelong Shen, Jingjing Liu and Jianfeng Gao	2644
<i>Semantically-Aligned Equation Generation for Solving and Reasoning Math Word Problems</i> Ting-Rui Chiang and Yun-Nung Chen	2656
<i>Iterative Search for Weakly Supervised Semantic Parsing</i> Pradeep Dasigi, Matt Gardner, Shikhar Murty, Luke Zettlemoyer and Eduard Hovy	2669
<i>Alignment over Heterogeneous Embeddings for Question Answering</i> Vikas Yadav, Steven Bethard and Mihai Surdeanu	2681
<i>Bridging the Gap: Attending to Discontinuity in Identification of Multiword Expressions</i> Omid Rohanian, Shiva Taslimipour, Samaneh Kouchaki, Le An Ha and Ruslan Mitkov	2692
<i>Incorporating Word Attention into Character-Based Word Segmentation</i> Shohei Higashiyama, Masao Utiyama, Eiichiro Sumita, Masao Ideuchi, Yoshiaki Oida, Yohei Sakamoto and Isaac Okada	2699
<i>VCWE: Visual Character-Enhanced Word Embeddings</i> Chi Sun, Xipeng Qiu and Xuanjing Huang	2710
<i>Subword Encoding in Lattice LSTM for Chinese Word Segmentation</i> Jie Yang, Yue Zhang and Shuailong Liang	2720
<i>Improving Cross-Domain Chinese Word Segmentation with Word Embeddings</i> Yuxiao Ye, Weikang Li, Yue Zhang, Likun Qiu and Jian Sun	2726
<i>Neural Semi-Markov Conditional Random Fields for Robust Character-Based Part-of-Speech Tagging</i> Apostolos Kemos, Heike Adel and Hinrich Schütze	2736
<i>Shrinking Japanese Morphological Analyzers With Neural Networks and Semi-supervised Learning</i> Arseny Tolmachev, Daisuke Kawahara and Sadao Kurohashi	2744

<i>Neural Constituency Parsing of Speech Transcripts</i> Paria Jamshid Lou, Yufei Wang and Mark Johnson	2756
<i>Acoustic-to-Word Models with Conversational Context Information</i> Suyoun Kim and Florian Metze	2766
<i>A Dynamic Speaker Model for Conversational Interactions</i> Hao Cheng, Hao Fang and Mari Ostendorf	2772
<i>Fluent Translations from Disfluent Speech in End-to-End Speech Translation</i> Elizabeth Salesky, Matthias Sperber and Alexander Waibel	2786
<i>Relation Classification Using Segment-Level Attention-based CNN and Dependency-based RNN</i> Van-Hien Tran, Van-Thuy Phi, Hiroyuki Shindo and Yuji Matsumoto	2793
<i>Document-Level Event Factuality Identification via Adversarial Neural Network</i> Zhong Qian, Peifeng Li, Qiaoming Zhu and Guodong Zhou	2799
<i>Distant Supervision Relation Extraction with Intra-Bag and Inter-Bag Attentions</i> Zhi-Xiu Ye and Zhen-Hua Ling	2810
<i>Ranking-Based Autoencoder for Extreme Multi-label Classification</i> Bingyu Wang, Li Chen, Wei Sun, Kechen Qin, Kefeng Li and Hui Zhou	2820
<i>Posterior-regularized REINFORCE for Instance Selection in Distant Supervision</i> Qi Zhang, Siliang Tang, Xiang Ren, Fei Wu, Shiliang Pu and Yueting Zhuang	2831
<i>Scalable Collapsed Inference for High-Dimensional Topic Models</i> Rashidul Islam and James Foulds	2836
<i>An Integrated Approach for Keyphrase Generation via Exploring the Power of Retrieval and Extraction</i> Wang Chen, Hou Pong Chan, Piji Li, Lidong Bing and Irwin King	2846
<i>Predicting Malware Attributes from Cybersecurity Texts</i> Arpita Roy, Youngja Park and Shimei Pan	2857
<i>Improving Distantly-supervised Entity Typing with Compact Latent Space Clustering</i> Bo Chen, Xiaotao Gu, Yufeng Hu, Siliang Tang, Guoping Hu, Yueting Zhuang and Xiang Ren	2862
<i>Modelling Instance-Level Annotator Reliability for Natural Language Labelling Tasks</i> Maolin Li, Arvid Fahlström Myrman, Tingting Mu and Sophia Ananiadou	2873
<i>Review-Driven Multi-Label Music Style Classification by Exploiting Style Correlations</i> Guangxiang Zhao, Jingjing Xu, Qi Zeng, Xuancheng Ren and Xu Sun	2884
<i>Fact Discovery from Knowledge Base via Facet Decomposition</i> Zihao Fu, Yankai Lin, Zhiyuan Liu and Wai Lam	2892
<i>A Richer-but-Smarter Shortest Dependency Path with Attentive Augmentation for Relation Extraction</i> Duy-Cat Can, Hoang-Quynh Le, Quang-Thuy Ha and Nigel Collier	2902
<i>Bidirectional Attentive Memory Networks for Question Answering over Knowledge Bases</i> Yu Chen, Lingfei Wu and Mohammed J Zaki	2913

<i>BoolQ: Exploring the Surprising Difficulty of Natural Yes/No Questions</i> Christopher Clark, Kenton Lee, Ming-Wei Chang, Tom Kwiatkowski, Michael Collins and Kristina Toutanova	2924
<i>Enhancing Key-Value Memory Neural Networks for Knowledge Based Question Answering</i> Kun Xu, Yuxuan Lai, Yansong Feng and Zhiguo Wang	2937
<i>Repurposing Entailment for Multi-Hop Question Answering Tasks</i> Harsh Trivedi, Heeyoung Kwon, Tushar Khot, Ashish Sabharwal and Niranjan Balasubramanian	2948
<i>GenderQuant: Quantifying Mention-Level Genderedness</i> , Nitya Parthasarathi and Sameer Singh	2959
<i>Analyzing Polarization in Social Media: Method and Application to Tweets on 21 Mass Shootings</i> Dorottya Demszky, Nikhil Garg, Rob Voigt, James Zou, Jesse Shapiro, Matthew Gentzkow and Dan Jurafsky	2970
<i>Learning to Decipher Hate Symbols</i> Jing Qian, Mai ElSherief, Elizabeth Belding and William Yang Wang	3006
<i>Long-tail Relation Extraction via Knowledge Graph Embeddings and Graph Convolution Networks</i> Ningyu Zhang, Shumin Deng, Zhanlin Sun, Guanying Wang, Xi Chen, Wei Zhang and Huajun Chen	3016
<i>GAN Driven Semi-distant Supervision for Relation Extraction</i> Pengshuai Li, Xinsong Zhang, Weijia Jia and Hai Zhao	3026
<i>A general framework for information extraction using dynamic span graphs</i> Yi Luan, Dave Wadden, Luheng He, Amy Shah, Mari Ostendorf and Hannaneh Hajishirzi ...	3036
<i>OpenCeres: When Open Information Extraction Meets the Semi-Structured Web</i> Colin Lockard, Prashant Shiralkar and Xin Luna Dong	3047
<i>Structured Minimally Supervised Learning for Neural Relation Extraction</i> Fan Bai and Alan Ritter	3057
<i>Neural Machine Translation of Text from Non-Native Speakers</i> Antonios Anastasopoulos, Alison Lui, Toan Q. Nguyen and David Chiang	3070
<i>Improving Domain Adaptation Translation with Domain Invariant and Specific Information</i> Shuhao Gu, Yang Feng and Qun Liu	3081
<i>Selective Attention for Context-aware Neural Machine Translation</i> Sameen Maruf, André F. T. Martins and Gholamreza Haffari	3092
<i>On Evaluation of Adversarial Perturbations for Sequence-to-Sequence Models</i> Paul Michel, Xian Li, Graham Neubig and Juan Pino	3103
<i>Accelerated Reinforcement Learning for Sentence Generation by Vocabulary Prediction</i> Kazuma Hashimoto and Yoshimasa Tsuruoka	3115
<i>Mitigating Uncertainty in Document Classification</i> Xuchao Zhang, Fanglan Chen, ChangTien Lu and Naren Ramakrishnan	3126

<i>Complexity-Weighted Loss and Diverse Reranking for Sentence Simplification</i> Reno Kriz, Joao Sedoc, Marianna Apidianaki, Carolina Zheng, Gaurav Kumar, Eleni Miltsakaki and Chris Callison-Burch	3137
<i>Predicting Helpful Posts in Open-Ended Discussion Forums: A Neural Architecture</i> Kishalay Halder, Min-Yen Kan and Kazunari Sugiyama	3148
<i>Text Classification with Few Examples using Controlled Generalization</i> Abhijit Mahabal, Jason Baldridge, Burcu Karagol Ayan, Vincent Perot and Dan Roth	3158
<i>Reinforcement Learning Based Text Style Transfer without Parallel Training Corpus</i> Hongyu Gong, Suma Bhat, Lingfei Wu, JinJun Xiong and Wen-mei Hwu	3168
<i>Adapting RNN Sequence Prediction Model to Multi-label Set Prediction</i> Kechen Qin, Cheng Li, Virgil Pavlu and Javed Aslam	3181
<i>Customizing Grapheme-to-Phoneme System for Non-Trivial Transcription Problems in Bangla Language</i> Sudipta Saha Shubha, Nafis Sadeq, Shafayat Ahmed, Md. Nahidul Islam, Muhammad Abdullah Adnan, Md. Yasin Ali Khan and Mohammad Zuberul Islam	3191
<i>Connecting Language and Knowledge with Heterogeneous Representations for Neural Relation Extrac- tion</i> Peng Xu and Denilson Barbosa	3201
<i>Segmentation-free compositional n-gram embedding</i> Geewook Kim, Kazuki Fukui and Hidetoshi Shimodaira	3207
<i>Exploiting Noisy Data in Distant Supervision Relation Classification</i> Kaijia Yang, Liang He, Xin-Yu Dai, Shujian Huang and Jiajun Chen	3216
<i>Misspelling Oblivious Word Embeddings</i> Aleksandra Piktus, Necati Bora Edizel, Piotr Bojanowski, Edouard Grave, Rui Ferreira and Fabrizio Silvestri	3226
<i>Learning Relational Representations by Analogy using Hierarchical Siamese Networks</i> Gaetano Rossiello, Alfio Gliozzo, Robert Farrell, Nicolas Fauceglia and Michael Glass	3235
<i>An Effective Label Noise Model for DNN Text Classification</i> Ishan Jindal, Daniel Pressel, Brian Lester and Matthew Nockleby	3246
<i>Understanding Learning Dynamics Of Language Models with SVCCA</i> Naomi Saphra and Adam Lopez	3257
<i>Using Large Corpus N-gram Statistics to Improve Recurrent Neural Language Models</i> Yiben Yang, Ji-Ping Wang and Doug Downey	3268
<i>Continual Learning for Sentence Representations Using Conceptors</i> Tianlin Liu, Lyle Ungar and Joao Sedoc	3274
<i>Relation Discovery with Out-of-Relation Knowledge Base as Supervision</i> Yan Liang, Xin Liu, Jianwen Zhang and Yangqiu Song	3280
<i>Corpora Generation for Grammatical Error Correction</i> Jared Lichtarge, Chris Alberti, Shankar Kumar, Noam Shazeer, Niki Parmar and Simon Tong	3291

<i>Structural Supervision Improves Learning of Non-Local Grammatical Dependencies</i> Ethan Wilcox, Peng Qian, Richard Futrell, Miguel Ballesteros and Roger Levy	3302
<i>Benchmarking Approximate Inference Methods for Neural Structured Prediction</i> Lifu Tu and Kevin Gimpel	3313
<i>Evaluating and Enhancing the Robustness of Dialogue Systems: A Case Study on a Negotiation Agent</i> Minhao Cheng, Wei Wei and Cho-Jui Hsieh	3325
<i>Investigating Robustness and Interpretability of Link Prediction via Adversarial Modifications</i> Pouya Pezeshkpour, Yifan Tian and Sameer Singh	3336
<i>Analysis Methods in Neural Language Processing: A Survey</i> Yonatan Belinkov and James Glass	3348
<i>Transferable Neural Projection Representations</i> Chinnadhurai Sankar, Sujith Ravi and Zornitsa Kozareva	3355
<i>Semantic Role Labeling with Associated Memory Network</i> Chaoyu Guan, Yuhao Cheng and Hai Zhao	3361
<i>Better, Faster, Stronger Sequence Tagging Constituent Parsers</i> David Vilares, Mostafa Abdou and Anders Søgaard	3372
<i>CAN-NER: Convolutional Attention Network for Chinese Named Entity Recognition</i> Yuying Zhu and Guoxin Wang	3384
<i>Decomposed Local Models for Coordinate Structure Parsing</i> Hiroki Teranishi, Hiroyuki Shindo and Yuji Matsumoto	3394
<i>Multi-Task Learning for Japanese Predicate Argument Structure Analysis</i> Hikaru Omori and Mamoru Komachi	3404
<i>Domain adaptation for part-of-speech tagging of noisy user-generated text</i> Luisa März, Dietrich Trautmann and Benjamin Roth	3415
<i>Neural Chinese Address Parsing</i> Hao Li, Wei Lu, Pengjun Xie and Linlin Li	3421
<i>Learning Hierarchical Discourse-level Structure for Fake News Detection</i> Hamid Karimi and Jiliang Tang	3432
<i>DiscoFuse: A Large-Scale Dataset for Discourse-Based Sentence Fusion</i> Mor Geva, Eric Malmi, Idan Szpektor and Jonathan Berant	3443
<i>Linguistically-Informed Specificity and Semantic Plausibility for Dialogue Generation</i> Wei-Jen Ko, Greg Durrett and Junyi Jessy Li	3456
<i>Learning to Describe Unknown Phrases with Local and Global Contexts</i> Shonosuke Ishiwatari, Hiroaki Hayashi, Naoki Yoshinaga, Graham Neubig, Shoetsu Sato, Masashi Toyoda and Masaru Kitsuregawa	3467
<i>Mining Discourse Markers for Unsupervised Sentence Representation Learning</i> Damien Sileo, Tim Van de Cruys, Camille Pradel and Philippe Muller	3477

<i>How Large a Vocabulary Does Text Classification Need? A Variational Approach to Vocabulary Selection</i> Wenhu Chen, Yu Su, Yilin Shen, Zhiyu Chen, Xifeng Yan and William Yang Wang	3487
<i>Subword-based Compact Reconstruction of Word Embeddings</i> Shota Sasaki, Jun Suzuki and Kentaro Inui	3498
<i>Bayesian Learning for Neural Dependency Parsing</i> Ehsan Shareghi, Yingzhen Li, Yi Zhu, Roi Reichart and Anna Korhonen	3509
<i>AutoSeM: Automatic Task Selection and Mixing in Multi-Task Learning</i> Han Guo, Ramakanth Pasunuru and Mohit Bansal	3520
<i>Studying the Inductive Biases of RNNs with Synthetic Variations of Natural Languages</i> Shauli Ravfogel, Yoav Goldberg and Tal Linzen	3532
<i>Attention is not Explanation</i> Sarthak Jain and Byron C. Wallace	3543
<i>Playing Text-Adventure Games with Graph-Based Deep Reinforcement Learning</i> Prithviraj Ammanabrolu and Mark Riedl	3557
<i>Information Aggregation for Multi-Head Attention with Routing-by-Agreement</i> Jian Li, Baosong Yang, Zi-Yi Dou, Xing Wang, Michael R. Lyu and Zhaopeng Tu	3566
<i>Context Dependent Semantic Parsing over Temporally Structured Data</i> Charles Chen and Razvan Bunescu	3576
<i>Structural Scaffolds for Citation Intent Classification in Scientific Publications</i> Arman Cohan, Waleed Ammar, Madeleine van Zuylen and Field Cady	3586
<i>pair2vec: Compositional Word-Pair Embeddings for Cross-Sentence Inference</i> Mandar Joshi, Eunsol Choi, Omer Levy, Daniel Weld and Luke Zettlemoyer	3597
<i>Submodular Optimization-based Diverse Paraphrasing and its Effectiveness in Data Augmentation</i> Ashutosh Kumar, Satwik Bhattamishra, Manik Bhandari and Partha Talukdar	3609
<i>Let's Make Your Request More Persuasive: Modeling Persuasive Strategies via Semi-Supervised Neural Nets on Crowdfunding Platforms</i> Diyi Yang, Jiaao Chen, Zichao Yang, Dan Jurafsky and Eduard Hovy	3620
<i>Recursive Routing Networks: Learning to Compose Modules for Language Understanding</i> Ignacio Cases, Clemens Rosenbaum, Matthew Riemer, Atticus Geiger, Tim Klinger, Alex Tamkin, Olivia Li, Sandhini Agarwal, Joshua D. Greene, Dan Jurafsky, Christopher Potts and Lauri Karttunen	3631
<i>Structural Neural Encoders for AMR-to-text Generation</i> Marco Damonte and Shay B. Cohen	3649
<i>Multilingual prediction of Alzheimer's disease through domain adaptation and concept-based language modelling</i> Kathleen C. Fraser, Nicklas Linz, Bai Li, Kristina Lundholm Fors, Frank Rudzicz, Alexandra Konig, Jan Alexandersson, Philippe Robert and Dimitrios Kokkinakis	3659
<i>Ranking and Selecting Multi-Hop Knowledge Paths to Better Predict Human Needs</i> Debjit Paul and Anette Frank	3671

<i>NLP Whack-A-Mole: Challenges in Cross-Domain Temporal Expression Extraction</i> Amy Olex, Luke Maffey and Bridget McInnes	3682
<i>Document-Level N-ary Relation Extraction with Multiscale Representation Learning</i> Robin Jia, Cliff Wong and Hoifung Poon	3693
<i>Inferring Which Medical Treatments Work from Reports of Clinical Trials</i> Eric Lehman, Jay DeYoung, Regina Barzilay and Byron C. Wallace	3705
<i>Decay-Function-Free Time-Aware Attention to Context and Speaker Indicator for Spoken Language Understanding</i> Jonggu Kim and Jong-Hyeok Lee	3718
<i>Dialogue Act Classification with Context-Aware Self-Attention</i> Vipul Raheja and Joel Tetreault	3727
<i>Affect-Driven Dialog Generation</i> Pierre Colombo, Wojciech Witon, Ashutosh Modi, James Kennedy and Mubbasir Kapadia ...	3734
<i>Multi-Level Memory for Task Oriented Dialogs</i> Revanth Gangi Reddy, Danish Contractor, Dinesh Raghu and Sachindra Joshi	3744
<i>Topic Spotting using Hierarchical Networks with Self Attention</i> Pooja Chitkara, Ashutosh Modi, Pravalika Avvaru, Sepehr Janghorbani and Mubbasir Kapadia	3755
<i>Top-Down Structurally-Constrained Neural Response Generation with Lexicalized Probabilistic Context-Free Grammar</i> Wenchao Du and Alan W. Black	3762
<i>What do Entity-Centric Models Learn? Insights from Entity Linking in Multi-Party Dialogue</i> Laura Aina, Carina Silberer, Ionut-Teodor Sorodoc, Matthijs Westera and Gemma Boleda ...	3772
<i>Continuous Learning for Large-scale Personalized Domain Classification</i> Han Li, Jihwan Lee, Sidharth Mudgal, Ruhi Sarikaya and Young-Bum Kim	3784
<i>Cross-lingual Transfer Learning for Multilingual Task Oriented Dialog</i> Sebastian Schuster, Sonal Gupta, Rushin Shah and Mike Lewis	3795
<i>Evaluating Coherence in Dialogue Systems using Entailment</i> Nouha Dziri, Ehsan Kamaloo, Kory Mathewson and Osmar Zaiane	3806
<i>On Knowledge distillation from complex networks for response prediction</i> Siddhartha Arora, Mitesh M. Khapra and Harish G. Ramaswamy	3813
<i>Cross-lingual Multi-Level Adversarial Transfer to Enhance Low-Resource Name Tagging</i> Lifu Huang, Heng Ji and Jonathan May	3823
<i>Unsupervised Extraction of Partial Translations for Neural Machine Translation</i> Benjamin Marie and Atsushi Fujita	3834
<i>Low-Resource Syntactic Transfer with Unsupervised Source Reordering</i> Mohammad Sadegh Rasooli and Michael Collins	3845
<i>Revisiting Adversarial Autoencoder for Unsupervised Word Translation with Cycle Consistency and Improved Training</i> Tasnim Mohiuddin and Shafiq Joty	3857

<i>Addressing word-order Divergence in Multilingual Neural Machine Translation for extremely Low Resource Languages</i>	
Rudra Murthy, Anoop Kunchukuttan and Pushpak Bhattacharyya	3868
<i>Massively Multilingual Neural Machine Translation</i>	
Roe Aharoni, Melvin Johnson and Orhan Firat	3874
<i>A Large-Scale Comparison of Historical Text Normalization Systems</i>	
Marcel Bollmann	3885
<i>Combining Discourse Markers and Cross-lingual Embeddings for Synonym–Antonym Classification</i>	
Michael Roth and Shyam Upadhyay	3899
<i>Context-Aware Cross-Lingual Mapping</i>	
Hanan Aldarmaki and Mona Diab	3906
<i>Polyglot Contextual Representations Improve Crosslingual Transfer</i>	
Phoebe Mulcaire, Jungo Kasai and Noah A. Smith	3912
<i>Typological Features for Multilingual Delexicalised Dependency Parsing</i>	
Manon Scholivet, Franck Dary, Alexis Nasr, Benoit Favre and Carlos Ramisch	3919
<i>Recommendations for Datasets for Source Code Summarization</i>	
Alexander LeClair and Collin McMillan	3931
<i>Question Answering as an Automatic Evaluation Metric for News Article Summarization</i>	
Matan Eyal, Tal Baumel and Michael Elhadad	3938
<i>Understanding the Behaviour of Neural Abstractive Summarizers using Contrastive Examples</i>	
Krtin Kumar and Jackie Chi Kit Cheung	3949
<i>Jointly Extracting and Compressing Documents with Summary State Representations</i>	
Afonso Mendes, Shashi Narayan, Sebastião Miranda, Zita Marinho, André F. T. Martins and Shay B. Cohen	3955
<i>News Article Teaser Tweets and How to Generate Them</i>	
Sanjeev Kumar Karn, Mark Buckley, Ulli Waltinger and Hinrich Schütze	3967
<i>Cross-referencing Using Fine-grained Topic Modeling</i>	
Jeffrey Lund, Piper Armstrong, Wilson Fearn, Stephen Cowley, Emily Hales and Kevin Seppi	3978
<i>Conversation Initiation by Diverse News Contents Introduction</i>	
Satoshi Akasaki and Nobuhiro Kaji	3988
<i>Positional Encoding to Control Output Sequence Length</i>	
Sho Takase and Naoaki Okazaki	3999
<i>The Lower The Simpler: Simplifying Hierarchical Recurrent Models</i>	
Chao Wang and Hui Jiang	4005
<i>Using Natural Language Relations between Answer Choices for Machine Comprehension</i>	
Rajkumar Pujari and Dan Goldwasser	4010
<i>Saliency Learning: Teaching the Model Where to Pay Attention</i>	
Reza Ghaeini, Xiaoli Fern, Hamed Shahbazi and Prasad Tadepalli	4016

<i>Understanding Dataset Design Choices for Multi-hop Reasoning</i> Jifan Chen and Greg Durrett	4026
<i>Neural Grammatical Error Correction with Finite State Transducers</i> Felix Stahlberg, Christopher Bryant and Bill Byrne	4033
<i>Convolutional Self-Attention Networks</i> Baosong Yang, Longyue Wang, Derek F. Wong, Lidia S. Chao and Zhaopeng Tu	4040
<i>Rethinking Complex Neural Network Architectures for Document Classification</i> Ashutosh Adhikari, Achyudh Ram, Raphael Tang and Jimmy Lin	4046
<i>Pre-trained language model representations for language generation</i> Sergey Edunov, Alexei Baevski and Michael Auli	4052
<i>Pragmatically Informative Text Generation</i> Sheng Shen, Daniel Fried, Jacob Andreas and Dan Klein	4060
<i>Stochastic Wasserstein Autoencoder for Probabilistic Sentence Generation</i> Hareesh Bahuleyan, Lili Mou, Hao Zhou and Olga Vechtomova	4068
<i>Benchmarking Hierarchical Script Knowledge</i> Yonatan Bisk, Jan Buys, Karl Pichotta and Yejin Choi	4077
<i>A large-scale study of the effects of word frequency and predictability in naturalistic reading</i> Cory Shain	4086
<i>Augmenting word2vec with latent Dirichlet allocation within a clinical application</i> Akshay Budhkar and Frank Rudzicz	4095
<i>On the Idiosyncrasies of the Mandarin Chinese Classifier System</i> Shijia Liu, Hongyuan Mei, Adina Williams and Ryan Cotterell	4100
<i>Joint Learning of Pre-Trained and Random Units for Domain Adaptation in Part-of-Speech Tagging</i> Sara Meftah, Youssef Tamaazousti, Nasredine Semmar, Hassane Essafi and Fatiha Sadat	4107
<i>Show Some Love to Your n-grams: A Bit of Progress and Stronger n-gram Language Modeling Baselines</i> Ehsan Shareghi, Daniela Gerz, Ivan Vulić and Anna Korhonen	4113
<i>Training Data Augmentation for Context-Sensitive Neural Lemmatizer Using Inflection Tables and Raw Text</i> Toms Bergmanis and Sharon Goldwater	4119
<i>A Structural Probe for Finding Syntax in Word Representations</i> John Hewitt and Christopher D. Manning	4129
<i>CNM: An Interpretable Complex-valued Network for Matching</i> Qiuchi Li, Benyou Wang and Massimo Melucci	4139
<i>CommonsenseQA: A Question Answering Challenge Targeting Commonsense Knowledge</i> Alon Talmor, Jonathan Herzig, Nicholas Lourie and Jonathan Berant	4149
<i>Probing the Need for Visual Context in Multimodal Machine Translation</i> Ozan Caglayan, Pranava Madhyastha, Lucia Specia and Loïc Barrault	4159

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding
Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova 4171

What's in a Name? Reducing Bias in Bios without Access to Protected Attributes
Alexey Romanov, Maria De-Arteaga, Hanna Wallach, Jennifer Chayes, Christian Borgs, Alexandra Chouldechova, Sahin Geyik, Krishnaram Kenthapadi, Anna Rumshisky and Adam Kalai 4187