# A Complete Bibliography of ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)

Nelson H. F. Beebe University of Utah Department of Mathematics, 110 LCB 155 S 1400 E RM 233 Salt Lake City, UT 84112-0090 USA

> Tel: +1 801 581 5254 FAX: +1 801 581 4148

E-mail: beebe@math.utah.edu, beebe@acm.org, beebe@computer.org (Internet)
WWW URL: http://www.math.utah.edu/~beebe/

08 July 2020 Version 1.06

# Title word cross-reference

K [30].  $\mu$  [172]. N [24, 26].  $S^3$  [195]. -Forcing [172]. -Gram [26, 24]. -Means [30]. -NET [195]. 15919 [188].

Abstractive [143]. Accented [6]. Acoustic [22]. Act [95, 18]. Action [26]. Adaptation [77, 32]. Additional [144]. Advanced [215]. Advances [139]. Adversarial [177, 213]. Agents [157]. Agglutinative [137]. Algorithm [15]. Alignment [29, 82]. All-Words [152]. Analysis [30, 127, 132, 106, 71, 103, 189, 181, 200, 165, 50, 109, 151, 170, 84, 56, 142, 159]. Analyzing [141]. Ancient [166]. Animate [107]. Animation [72].

Annotate [111]. Annotated [45, 57, 152]. Annotation [14, 129, 101, 138, 183]. Answering [185, 204, 59]. Any [183, 220]. Application [4, 97]. Applications [139, 168]. Approach [189, 48, 155, 116, 4, 162, 114, 36, 86, 57, 76, 59, 187, 82, 17, 163]. Approaches [2]. Arabic [182, 30, 73, 81, 132, 117, 7, 139, 71, 189, 135, 153, 31, 122, 217, 192, 99, 89, 59, 27, 16, 95, 70]. Arabizi [192]. AROMA [73]. Array [46]. Artificial [79]. Asian [144, 207]. Aspect [127, 159]. Aspect-Based [127, 159]. Aspects [141]. Assamese [84]. assisted [173]. Attention [155, 158, 219]. Attentive [189]. Attribution [117, 76]. Authorship [117, 76, 196]. Autoencoders [172]. Automatic [101, 29, 114, 140, 173, 115]. Automatically [81]. Auxiliary [68, 179]. Aware [75].

Balancing [197]. Bangla [85, 48, 97]. Base [92, 204, 131]. Based [130, 127, 132, 65, 14, 155, 148, 120, 122, 88, 22, 4, 80, 53, 124, 126, 128, 195, 59, 49, 201, 115, 214, 104, 68, 82, 20, 56, 108, 167, 218, 17, 183, 159, 37, 156, 118, 198, 178, 182, 146, 165, 190, 157, 163]. Basic [87]. Bengali [25, 76]. BenLem [25]. Bigram [8]. Bilingual [63, 148, 11, 86, 57, 6, 173, 24, 108, 17, 28]. Bisect [30]. Boosted [58]. Boosting [52]. Bootstrapping [48]. Boundary [94]. Bringing [209]. Broadcast [75, 205]. Broadened [5]. Building [81]. Burmese [41, 200, 165, 118].

Canonical [106]. Carrier [180]. Case [127, 23, 197, 55, 78]. Catalan [155]. Categorial [3]. Categories [66]. Causal [27]. Causality [157]. Cause [83]. CCG [1]. CDT [74]. CDT-Styled [74]. Cells [208]. Chain [162]. Chain-to-chain [162]. Challenges [139, 135]. Channel [153]. Character [48, 29, 107, 16, 91]. Characters [44, 97]. Check [35, 36, 69]. Checker [37]. Checking [34]. Chief [207]. Children [190]. Chinese [35, 83, 23, 14, 155, 148, 105, 38, 100, 92, 138, 74, 162, 34, 12, 29, 61, 67, 36, 53, 166, 60, 91, 167, 143, 37, 163, 69]. Chinese-to-Spanish [14]. Chunk [33]. CLASENTI [109]. Class [109]. Class-Specific [109]. Classes [7]. Classical [135]. Classification [211, 146, 153, 50, 212, 44, 190, 191, 22, 119, 79, 147]. Classify [98]. Clean [173]. Cleaning [85]. Close [77]. Clustering [30, 128]. Co [134]. Co-occurrence [134]. Coarse [158]. Coarse-to-fine [158]. Code [210, 141]. Code-mixed [210]. Code-Switching [141]. Coherent [19]. Collaborative [163]. Collective [20]. Combination [14]. Combinatory [3]. Combine [164]. Combining [193]. Commentary [39, 5]. Communities [113]. Compact [46]. Comparable [171, 23]. Comparative [50, 142, 218]. Comparison [111, 64]. Complex [19]. Components [64]. Composite [217]. Composition [64]. Comprehension [195]. Comprehensive [139]. Conditional [22, 10]. Conducting [206]. Confusion [53]. Connectivity [21]. Consistency [143]. Constituent [26]. Constraint [86, 17]. Construct [57]. Constructing [101, 19]. Content [70]. Context [216, 206, 54, 159, 18]. Context-Dependent [216]. Context-Free [54, 18]. Contextualized [169]. Continuous [24]. Continuous-Space [24]. Contribution [89]. Conversation [179]. Conversational [178]. Converting [145, 24]. Convolutional [153, 147]. ConvS2S [186]. Copying [96]. Corpora [81, 171, 23, 111, 68, 1]. Corpus [135, 83, 184, 168, 57, 152, 173, 68, 183]. Corpus-Based [68]. Correcting [38]. Correction [216, 120, 174]. Correlation [106]. Counterparts [77]. Coverage [127, 1]. Covering [139]. Creole [161]. CRFs [118]. Critical [64]. Cross [127, 113, 31, 13]. Cross-Language [113, 31, 13]. Cross-Linguality [127]. Cues [174].

Data [93, 193]. Database [40]. Dataset [45, 166, 49]. De-Training [186]. Decoding [38]. Decomposition [28]. Deep [73, 71, 189, 185, 168, 169, 98, 198]. Delexicalized [164]. Dependencies [161]. Dependency [65, 145, 206, 105, 136, 3, 169, 126, 51, 115, 1, 218, 159, 18, 156]. Dependency-Based [218]. Dependent [216]. Derived [113]. Description [14]. Devanagari [15]. Developed [14]. Development [84]. Dharma [187]. Diacritic [122]. Diacritic-Based [122]. Diacritics [140]. Dialect [182, 132, 140, 192]. Dialects [81]. Dialogues [151]. Dictionaries [70]. Dictionary [11, 86, 17]. Different [141]. Disambiguation [21, 152, 112]. Discourse [146, 88, 138, 74, 67, 59, 98]. Discourse-Based [59]. Distortions [16]. Distributed [132]. Distribution [143]. Document [62, 4, 55, 119, 110]. Document-Level [119, 110]. Documents [30, 132, 7, 220, 196]. Domain [110, 178]. Domain-specific [110]. Double [46, 180]. Double-Array [46]. Double-Carrier [180]. Dual [28, 118]. Dual-Layer [118]. Dynamic [204].

Editor [207]. Editor-in-Chief [207]. Editorial [39, 5, 207]. EDU [4]. EDU-Based [4]. Effective [71, 67, 124, 154]. Effectively [116]. Efficient [120, 24, 194]. Embedding [182, 127, 153, 100, 108, 180]. Embeddings [210, 66, 169, 134, 198]. Emotion [83, 212]. Emotional [179]. EmotionNet [212]. Empirical [91]. End [96, 74]. End-to-End [96, 74]. Endangered [42]. English [33, 185, 197, 67, 57, 141, 142]. Enhanced [205, 180, 158]. Enhancing [26]. Enriched [71, 54]. Entities [102]. Entity [211, 58, 66, 168, 111, 219, 79, 57, 187, 110, 159]. **Equivalence** [7]. **Era** [207]. Errors [38]. Establishment [92]. Estimation [80, 160, 197]. Estimator [80]. Evaluating [70]. Evaluation [29, 114, 213, 156]. Event [61]. Expanding [90]. Expansion [62, 115]. Experience [157]. Experience-based [157]. Experiment [70]. Experimental [176]. Explicitly [175]. Exploiting [90]. Exploring [91]. Expression [2]. Extensive [117]. Extra [33]. Extra-Chunk [33]. Extracting [171, 217, 27, 183]. Extraction [7, 23, 92, 15, 61, 11, 219, 124, 20, 208]. Extractive [75, 205, 142].

Facebook [132]. Families [86]. Farsi [52]. Fast [46]. Feasible [129]. Feature [123, 151, 15]. Features [48, 148, 44, 190, 22, 97, 47]. Feedback [194]. Fields [22, 10]. Fields-Based [22]. Filtered [184]. Filtering [163]. Filtering-based [163]. fine [158]. Fingerspelled [72]. Flexible [129]. Fluent [215]. Focus [5]. Forcing [172]. Form [209, 85]. Four [45]. Four-Tier [45]. Fragment [23]. Framework [62, 35, 185, 109, 75, 196]. Free [85, 54, 18]. Free-form [85]. Fusion [193]. Fuzzy [21].

Generalities [90]. Generalized [86, 47]. Generating [53]. Generation [168, 172, 6, 201, 214, 134, 179, 178]. Genesis [161]. Genitive [217]. Gnani [220]. Goals [19]. Gram [26, 24]. Grammar [217, 3, 214]. Grammars [54, 18]. Graph [171, 21, 108]. Graph-Based [108]. Graphs [18]. guided [191]. Guideline [200]. Gujarati [44]. Gurmukhi [49, 133].

Handling [124]. Handwriting [123, 16]. Handwritten [85, 48, 45, 147, 97, 16, 49, 133, 8]. Heritage [99]. Hidden [22]. Hierarchical [98]. High [150]. High-Resource [150]. Hindi [65, 212, 185, 21, 120, 55, 2, 141, 142]. HindiWordNet [131]. Hinge [218]. HTK [84]. Human [14, 116, 16]. Hybrid [14, 15, 36, 126, 69].

Identification [202, 107, 215, 196, 18]. Identifying [141]. Igbo [87, 154]. Image [45, 40, 42]. Imperceptibly [116]. Implementation [46, 87]. Implicit [67, 98]. Importance [174]. Important [103]. Improved [88, 8, 143]. Improves [184]. Improving [127, 106, 103, 65, 210, 3, 54, 203, 121, 51, 16, 82, 144]. Inclusive [133]. Incorporating [100, 119, 149]. Index [37]. Indian [113, 190, 174, 78, 214, 68]. Indic [123, 176, 193]. Inducing [63, 1]. Induction [93, 86, 68, 17]. Inference [206, 28]. Information [146, 171, 113, 31, 131, 205, 193, 52, 149, 159, 32]. Injection [78]. Innovations [207]. Input [116]. Integrate [116]. Integrated [23]. Integrating [94, 1]. Intelligent [157]. Inter [33]. Inter- [33]. Interpolated [197]. Interpolation [203]. Intra [33]. Intra- [33]. Introduction [34]. Inverted [37]. Iraqi [182, 132]. Isarn [187]. ISO [188]. Issue [34].

Japanese [23, 33, 151, 1]. Japanese-to-English [33]. Joint [137, 129, 219]. Journal [39, 5].

Kannada [220]. Kernel [106]. Key [135]. Keyword [7, 190]. Keyword-based [190]. Keywords [178]. Kit [87]. Knowledge [100, 92, 131, 204, 51]. Knowledge-base [131]. Korean [96, 136, 191, 197, 10, 126, 107, 201]. Kurdish [130].

**Labeling** [199, 28, 47]. **Labels** [94]. **Language** [73, 103, 113, 135, 66, 206, 31, 13, 105, 120, 191, 220, 75, 205, 170, 199, 86, 46,

87, 72, 78, 84, 215, 214, 8, 68, 213, 70, 24, 161, 32, 207]. Languages [127, 62, 113, 137, 176, 164, 190, 150, 202, 121, 77, 174, 68, 144, 134, 207]. Languages-An [176]. Large [166, 204]. Large-Scale [204]. Lattice [38]. Layer [186, 118]. Layer-Wise [186]. Learning [73, 189, 102, 120, 168, 67, 157, 150, 121, 76, 98, 161, 47, 198, 146]. Lemmatizer [25]. Less [127]. Less-Resourced [127]. Level [44, 119, 110]. Leveraging [98, 144]. Lexical [131]. Lexicon [103, 181, 63, 212]. Lexicons [90]. Linear [203]. Linguality [127]. Linguistic [190, 27, 82]. Linguistic-Relationships-Based [82]. Lipi [220]. List [37]. ListMLE [114]. Literary [76]. LMF [70]. Loanword [202]. Local [206]. Logic [20]. Logs [51]. Loss [218]. Low [73, 176, 66, 164, 44, 184, 125, 150, 93, 202, 124, 121, 86, 77, 174, 144, 194, 134, 207]. Low-Level [44]. Low-Resource [176, 66, 125, 150, 93, 202, 124, 121, 86, 77, 174, 144, 194, 207, 164, 184]. LSTMs [126].

Machine [146, 14, 155, 33, 148, 13, 175, 116, 184, 177, 80, 125, 54, 114, 203, 93, 124, 195, 173, 60, 144, 24, 108, 158, 156, 194, 186, 209, 9]. Machine-assisted [173]. Machine-learning [146]. Mandarin [198]. Manipuri [147]. Manual [101]. Markov [20]. Markovian [56]. Match [167]. Matching [171, 122, 195]. Math [149]. Mayek [147]. Means [30]. Measurement [100]. Measurements [128]. Measures [21]. Mechanism [96, 219, 37, 194]. Mechanisms [155]. Media [143, 209]. Memory [159, 208]. Message [115]. Message-Response [115]. Method [146, 171, 116, 53, 98, 173]. Methods [111, 213]. Metric [156]. Microblogs [83]. Minimal [202, 49]. Minimally [61]. Mining [73, 139]. mixed [210]. Model [73, 189, 153, 88, 12, 219, 46, 94, 6, 115, 56, 64, 156, 69, 26]. Modeling [65, 175, 29, 67, 75, 205, 16, 198]. **Models** [139, 71, 11, 8, 24, 32, 179]. Modern [182, 166]. Monolingual [29, 93]. Morpheme [10, 201]. Morphemes [126]. Morphological [176, 165, 99, 52]. Morphologically [71]. Morphology [78, 56]. Multi [127, 153, 4, 160, 119, 150, 149, 159]. Multi-Channel [153]. Multi-Document [4]. Multi-Entity [159]. Multi-Level [119]. Multi-Linguality [127]. Multi-Round [150]. Multi-Sentence [149]. Multi-task [160]. Multidisciplinary [45]. Multilingual [62, 113, 11, 121, 128]. Multiple [83, 61, 150, 1]. Multiple-User [83]. Multitask [136]. Multiview [67]. Multiword [63]. Mutual [98]. Myanmar [41, 200, 165].

Named [211, 58, 66, 168, 111, 79, 57, 187, 110]. Names [217]. Native [215]. Natural [206, 183, 207]. NER [121]. NET [195]. Network [153, 185, 136, 79, 147]. Networks [125, 195, 20, 32, 198]. NeuMorph [176]. Neural [176, 155, 153, 185, 175, 184, 177, 88, 80, 160, 125, 203, 79, 147, 52, 149, 213, 218, 158, 194, 32, 198, 179]. News [75, 205]. NLP [87]. NMT [150]. Non [215, 209]. Non-Native [215]. Non-Standard [209]. Normalization [209]. Nouns [107]. NOVA [129]. Novel [116, 107]. Number [72]. Nyushu

[42].

occurrence [134]. OCR [40, 220]. Offline [193]. Online [85, 123, 193, 97, 49, 133, 8]. Ontology [182]. OOV [12, 124]. Open [92, 178]. Open-Domain [178]. Opinion [73, 139]. Optimization [110]. Optimizing [114]. Order [178]. Order-Sensitive [178]. Ordering [33]. Outline [208]. Output [143].

Pair [206, 115]. Pairwise [50]. Parallel [81, 171, 23, 63, 184]. Paraphrase [90]. Parenthetical [20]. Parser [164, 13, 105, 74, 149]. Parsing [65, 88, 136, 3, 54, 169, 126, 51, 167, 161, 156, 26]. Part [129, 165, 96, 99, 154, 201]. Part-of-Speech [129, 96, 99, 154, 165, 201]. Path [203]. Patterns [27, 183]. Pause [124]. Pause-Based [124]. Performance [121]. Persian [103, 145, 181, 199]. Perspective [139, 78]. Perspectives [39, 5]. Phenomena [65]. Phonetic [188]. Phonetics [180]. Phrase [145, 93, 124, 94, 64]. Phrase-Boundary [94]. Picture [43]. Pivot [17]. Pivot-Based [17]. Pivoting [155]. Pointer [136]. Polarity [181, 183]. Poor [170]. POS [210, 10, 52, 161, 158, 112, 137]. Position [75]. Position-Aware [75]. Posts [117]. Practical [70]. Pre [33]. Pre-Ordering [33]. Prediction [80, 197]. Predictor [80]. Preference [119]. Preordering [13]. Principles [217]. Printed [44, 40, 220]. Prior [100]. Probabilistic [35, 18]. Problem [12]. Problems [177, 149]. Processing [131, 174, 42, 207]. Project [87]. Projection [13]. Pronoun [162, 163]. Propagation [160]. Proximity [205]. Pruning [24]. Pseudo [184]. Pseudo-parallel [184]. Punjabi [188, 211].

**QA** [115]. **Quality** [80, 160, 197]. **Query** [62, 51]. **Question** [185, 204, 59, 208]. **Question-Specific** [208].

Random [22, 10]. Ranking [36]. Re [60, 186]. Re-Segmentation [60]. Re-Training [186]. Reading [195]. Recognition [211, 123, 58, 66, 168, 67, 219, 79, 193, 2, 84, 16, 6, 95, 133, 187, 8, 110]. Recognize [97]. Recommend [102]. Rectification [188]. Recurrent [125, 172]. Recursive [73, 71]. Reduce [26]. Reevaluation [8]. Refining [103]. Regularizing [143]. Related [135, 102]. Relatedness [205, 128]. Relation [92, 67, 219, 107]. Relations [27, 98]. Relationship [91]. Relationships [82]. Relative [123]. Relevant [65]. Reordering [191]. Representation [67, 64, 91, 218]. Representations [132, 106, 138, 126]. Reread [194]. Rescoring [37]. Research [135, 45]. Resolution [162, 163]. Resource [73, 62, 135, 176, 66, 125, 150, 93, 170, 202, 124, 121, 86, 87, 77, 174, 144, 194, 134, 207, 164, 184]. Resource-Scarce [62]. Resourced [127]. Resources [139, 144, 1]. Response [115, 178]. Restoration [140]. Retrieval [113, 31, 115]. Retrieval-Based [115]. Review [31]. Robust [197, 213]. Role [146, 25, 199, 78, 28, 47]. Roman [188, 170]. Round [150]. Rule [130, 146, 14, 104]. Rule-Based [130, 14, 104, 146].

Scalable [212, 196]. Scale [204]. Scarce [62]. Script [45, 15, 220, 192, 42]. Search [102, 43, 19]. Seed [53]. Seed-Based [53]. Segmentation [48, 41, 12, 99, 10, 89, 187, 60, 118]. **Selection** [115, 134, 198]. **Self** [219, 195]. Self-attention [219]. Self-Matching [195]. Semantic [54, 199, 107, 115, 149, 167, 143, 28, 47, 18, 9]. **Semantics** [55, 89, 98]. Semisupervised [56]. Sense [21, 152]. Sensitive [178]. Sentence [23, 205, 195, 149, 91]. **SentiFars** [181]. **Sentiment** [127, 132, 146, 71, 103, 189, 153, 181, 212, 109, 151, 119, 170, 159]. Sequence [216]. Sequence-to-Sequence [216]. Sequences [63]. Sequential [213]. Serendipity [102]. Set [151, 49]. Sets [53]. Shahmukhi [211]. Shallow [94]. Shift [26]. Shifting [183]. Short [63, 203, 128]. Short-Path [203]. Sign [191, 72, 214]. Signal [174]. Signs [72]. Similarity [100]. Sindhi [40]. Singlish [161]. Sinhala [72]. Sliding [167]. SMT [78, 32]. Social [143, 209]. Solving [12]. Space [106, 24, 9]. Spanish [14]. Spatio [193]. Spatio-temporal [193]. Special [34]. Specific [109, 208, 110]. Speech [81, 129, 96, 75, 99, 154, 84, 6, 95, 18, 198, 165, 201]. Spell [34]. Spelling [216, 35, 38, 120, 36, 37, 69]. Spoken [105]. Spontaneous [105]. SRU [195]. SRU-Based [195]. Stack [160]. Standard [182, 209]. Standardized [70]. State [39]. Statistical [14, 33, 148, 13, 54, 93, 187, 144, 24, 108]. Stemmer [104]. Stemming [137, 89]. Step [88]. Story [190]. Strategy [8]. String [167]. Stroke [123, 44]. Strokes [49]. Structural [48, 97]. Structurally [218]. Structure [62, 145]. Structures [83, 46]. Study [127, 117, 176, 23, 197, 55, 78, 133, 64]. **Styled** [74]. **Stylometric** [50, 196]. StyloThai [196]. Sub [123]. Sub-Stroke-Wise [123]. Subtask [19]. Subword [189]. Suggested [43]. Summaries [55]. Summarization [4, 75, 205, 142, 143]. Sundanese [104]. Supertags [3]. Supervised [48, 120, 61, 76, 112, 179]. **Supervision** [202]. **Supplementary** [151]. Survey [139, 138]. Switching [141]. Syllable [165]. Syllable-based [165]. Syllables [126]. Synchronous [54]. Syntactic [13, 94]. Syntactically [65, 191]. Syntactically-guided [191]. Syntax [148, 167]. Syntax-Based [148]. Synthesis [198]. System [130, 139, 14, 129, 120, 204, 79, 199, 214]. Systems [124, 84, 115, 178].

Table [93]. Tag [56, 158]. Tag-enhanced [158]. Tag-Transition-Based [56]. Tagger [154, 52]. Tagging [210, 137, 176, 129, 165, 96, 99, 121, 10, 201, 213, 161, 112]. TALLIP [39, 5]. Tamil [8]. Target [13, 80]. Target-Language [13]. task [160]. Tasks [19]. Technique [30, 146]. Techniques [14, 95]. Telugu [3]. Tempo [131]. Tempo-HindiWordNet [131]. Temporal [131, 2, 193]. Term [7]. Text [130, 85, 45, 212, 40, 172, 193, 128, 142, 218, 143, 209]. Texts [55, 76]. Thai [22, 4, 196, 213]. Tier [45]. Tokenization [129, 165]. Tone [22]. Tools [139]. Topic [11, 89, 32]. Topological [97]. Topology [182]. Topology-based [182]. Training [58, 177, 166, 172, 186]. Transcription [174]. Transfer [164, 150, 161]. Transform [164]. Transformation [6]. Transition

[65, 88, 126, 56]. Transition-Based [65, 88, 126]. Translating [77]. Translation [14, 155, 33, 148, 13, 175, 116, 184, 177, 191, 80, 197, 54, 29, 114, 203, 166, 93, 124, 94, 68, 60, 144, 20, 24, 108, 158, 156, 194, 186, 9, 42]. Translations [175]. Translator [50]. Translators [116]. Transliteration [188, 130, 125, 192]. Tree [148]. Tree-to-Tree [148]. Treebank [71, 200]. Tri [58]. Tri-Training [58]. Tunisian [140, 192]. Turkish [216, 101, 43]. Twitter [117, 141]. Two [88]. Two-Step [88].

Understanding [55]. Unified [12]. Uniformly [197]. Unit [203, 198]. Unit-Selection-Based [198]. Universal [169, 161]. Unknown [177]. Unsupervised [137, 51]. Updating [204]. Urdu [146, 65, 45, 168, 79, 170, 152]. Usage [135]. User [83, 119]. Users [102]. Using [30, 113, 210, 48, 14, 101, 44, 190, 21, 96, 191, 217, 125, 150, 79, 93, 126, 147, 46, 52, 27, 84, 187, 68, 20, 18, 32, 198, 7, 13, 3, 6, 115]. Utilizing [62, 107]. UTTAM [120]. Uyghur [56].

Variational [172]. Vec [182]. Vector [106, 9]. Vectors [113]. Verification [6]. Versatile [220]. Via [106, 58, 13, 121, 98, 28, 180]. Vietnamese [148, 57, 60]. Views [61]. Visualizations [139]. Vocabulary [77]. VoIP [81].

Wasf [182]. Wasf-Vec [182]. Weakly [112]. Web [58, 102, 43, 20]. Web-Based [20]. Weight [134]. Wide [1]. Wide-Coverage [1]. Wikipedia [23, 66, 128, 68]. Wikipedia-Based [128]. Wise [123, 186]. without [112]. Word [182, 127, 106, 113, 66, 41, 206, 175, 38, 100, 21, 177, 191, 80, 12, 169, 193, 126, 152, 187, 60, 82, 91, 108, 134, 118, 180]. Word-Character [91]. Word-Pair-Dependency [206]. WordNet [101, 21]. Words [103, 113, 122, 72, 152, 49, 133, 8]. Writers [215]. Writing [180]. WSD [25].

**Zero** [162, 163].

# References

Uematsu:2015:IMD

[1] Sumire Uematsu, Takuya Matsuzaki, Hiroki Hanaoka, Yusuke Miyao, and Hideki Mima. Integrating multiple dependency corpora for inducing wide-coverage Japanese CCG resources. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(1): 1:1–1:??, January 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Ramrakhiyani:2015:ATE

[2] Nitin Ramrakhiyani and Prasenjit Majumder. Approaches to temporal expression recognition in Hindi. ACM Transactions on Asian and

Low-Resource Language Information Processing (TALLIP), 14(1):2:1–2:??, January 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Kumari:2015:ITD

[3] B. Venkata Seshu Kumari and Ramisetty Rajeshwara Rao. Improving Telugu dependency parsing using combinatory categorial grammar supertags. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(1):3:1–3:??, January 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Ketui:2015:EBA

[4] Nongnuch Ketui, Thanaruk Theeramunkong, and Chutamanee Onsuwan. An EDU-based approach for Thai multi-document summarization and its application. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(1):4:1–4:??, January 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Sproat:2015:TPE

[5] Richard Sproat. TALLIP perspectives: Editorial commentary: The broadened focus of the journal. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 14(1):5:1–5:??, January 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Shen:2015:MGA

[6] Han ping Shen, Chung hsien Wu, and Pei shan Tsai. Model generation of accented speech using model transformation and verification for bilingual speech recognition. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(2):6:1–6:??, March 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Awajan:2015:KEA

[7] Arafat Awajan. Keyword extraction from Arabic documents using term equivalence classes. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(2):7:1–7:??, March 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Sundaram:2015:BLM

[8] Suresh Sundaram and A. G. Ramakrishnan. Bigram language models and reevaluation strategy for improved recognition of online handwritten Tamil words. ACM Transactions on Asian and Low-Resource Language

Information Processing (TALLIP), 14(2):8:1–8:??, March 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Zhang:2015:TMT

[9] Jiajun Zhang, Shujie Liu, Mu Li, Ming Zhou, and Chengqing Zong. Towards machine translation in semantic vector space. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(2):9:1–9:??, March 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Na:2015:CRF

[10] Seung-Hoon Na. Conditional random fields for Korean morpheme segmentation and POS tagging. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(3):10:1–10:??, June 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Liu:2015:MTM

[11] Xiaodong Liu, Kevin Duh, and Yuji Matsumoto. Multilingual topic models for bilingual dictionary extraction. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 14(3):11:1–11:??, June 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Li:2015:UMS

[12] Xiaoqing Li, Chengqing Zong, and Keh yih Su. A unified model for solving the OOV problem of Chinese word segmentation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 14 (3):12:1–12:??, June 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Goto:2015:PUT

[13] Isao Goto, Masao Utiyama, Eiichiro Sumita, and Sadao Kurohashi. Preordering using a target-language parser via cross-language syntactic projection for statistical machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(3): 13:1–13:??, June 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Costa-Jussa:2016:DCS

[14] Marta R. Costa-Jussà and Jordi Centelles. Description of the Chinese-to-Spanish rule-based machine translation system developed using a hybrid combination of human annotation and statistical techniques. ACM

Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(1):1:1–1:??, January 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Khanduja:2016:HFE

[15] Deepti Khanduja, Neeta Nain, and Subhash Panwar. A hybrid feature extraction algorithm for Devanagari script. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(1): 2:1–2:??, January 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Shatnawi:2016:IHA

[16] Maad Shatnawi and Sherief Abdallah. Improving handwritten Arabic character recognition by modeling human handwriting distortions. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(1):3:1–3:??, January 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Wushouer:2016:CAP

[17] Mairidan Wushouer, Donghui Lin, Toru Ishida, and Katsutoshi Hirayama. A constraint approach to pivot-based bilingual dictionary induction. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(1):4:1–4:??, January 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Yeh:2016:SAI

[18] Jui-Feng Yeh. Speech act identification using semantic dependency graphs with probabilistic context-free grammars. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(1): 5:1–5:??, January 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Wang:2016:CCSa

[19] Ting-Xuan Wang and Wen-Hsiang Lu. Constructing complex search tasks with coherent subtask search goals. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(2): 6:1-6:??, February 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Tsai:2016:CWB

[20] Richard Tzong-Han Tsai. Collective Web-based parenthetical translation extraction using Markov logic networks. ACM Transactions on Asian

and Low-Resource Language Information Processing (TALLIP), 15(2): 7:1–7:??, February 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Jain:2016:FHW

[21] Amita Jain and D. K. Lobiyal. Fuzzy Hindi WordNet and word sense disambiguation using fuzzy graph connectivity measures. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(2):8:1–8:??, February 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Kertkeidkachorn:2016:AFH

[22] Natthawut Kertkeidkachorn, Proadpran Punyabukkana, and Atiwong Suchato. Acoustic features for hidden conditional random fields-based Thai tone classification. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(2):9:1–9:??, February 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Chu:2016:IPS

[23] Chenhui Chu, Toshiaki Nakazawa, and Sadao Kurohashi. Integrated parallel sentence and fragment extraction from comparable corpora: a case study on Chinese–Japanese Wikipedia. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(2):10:1–10:??, February 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Wang:2016:CCSb

[24] Rui Wang, Masao Utiyama, Isao Goto, Eiichiro Sumita, Hai Zhao, and Bao-Liang Lu. Converting continuous-space language models into N-gram language models with efficient bilingual pruning for statistical machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):11:1-11:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Chakrabarty:2016:BBL

[25] Abhisek Chakrabarty and Utpal Garain. BenLem (a Bengali lemmatizer) and its role in WSD. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):12:1–12:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Zhou:2016:ESR

[26] Hao Zhou, Shujian Huang, Junsheng Zhou, Yue Zhang, Huadong Chen, Xinyu Dai, Chuan Cheng, and Jiajun Chen. Enhancing shift-reduce

constituent parsing with action N-gram model. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15 (3):13:1–13:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Sadek:2016:EAC

[27] Jawad Sadek and Farid Meziane. Extracting Arabic causal relations using linguistic patterns. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):14:1–14:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Yang:2016:BSR

[28] Haitong Yang, Yu Zhou, and Chengqing Zong. Bilingual semantic role labeling inference via dual decomposition. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):15:1–15:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

**Li:2016:MMC** 

[29] Maoxi Li, Mingwen Wang, Hanxi Li, and Fan Xu. Modeling monolingual character alignment for automatic evaluation of Chinese translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(3):16:1–16:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

 ${\bf Abuaiadah: 2016: UBM}$ 

[30] Diab Abuaiadah. Using bisect K-means clustering technique in the analysis of Arabic documents. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):17:1–17:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Elayeb:2016:ACL

[31] Bilel Elayeb and Ibrahim Bounhas. Arabic cross-language information retrieval: a review. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(3):18:1–18:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Zhao:2016:ALM

[32] Yinggong Zhao, Shujian Huang, Xin-Yu Dai, and Jiajun Chen. Adaptation of language models for SMT using neural networks with topic information. ACM Transactions on Asian and Low-Resource Language In-

formation Processing (TALLIP), 15(3):19:1–19:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Ding:2016:IIE

[33] Chenchen Ding, Keisuke Sakanushi, Hirona Touji, and Mikio Yamamoto. Inter-, intra-, and extra-chunk pre-ordering for statistical Japanese-to-English machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(3):20:1–20:??, March 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Lee:2015:ISI

[34] Lung-Hao Lee, Gina-Anne Levow, Shih-Hung Wu, and Chao-Lin Liu. Introduction to the special issue on Chinese spell checking. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 14(4):14:1–14:??, October 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

Chen:2015:PFC

[35] Kuan-Yu Chen, Hsin-Min Wang, and Hsin-Hsi Chen. A probabilistic framework for Chinese spelling check. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(4):15:1–15:??, October 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

Liu:2015:HRA

[36] Xiaodong Liu, Fei Cheng, Kevin Duh, and Yuji Matsumoto. A hybrid ranking approach to Chinese spelling check. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(4): 16:1–16:??, October 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

Yeh:2015:CSC

[37] Jui-Feng Yeh, Wen-Yi Chen, and Mao-Chuan Su. Chinese spelling checker based on an inverted index list with a rescoring mechanism. *ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP)*, 14(4):17:1–17:??, October 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

Hsieh:2015:CCS

[38] Yu-Ming Hsieh, Ming-Hong Bai, Shu-Ling Huang, and Keh-Jiann Chen. Correcting Chinese spelling errors with word lattice decoding. ACM

Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(4):18:1–18:??, October 2015. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

# Anonymous:2015:TPE

[39] Anonymous. TALLIP perspectives: Editorial commentary: The state of the journal. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 14(4):19:1–19:??, October 2015. CO-DEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). Special issue on Chinese spell checking.

# Hakro:2016:PTI

[40] Dil Nawaz Hakro and Abdullah Zawawi Talib. Printed text image database for Sindhi OCR. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(4):21:1–21:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Ding:2016:WSB

[41] Chenchen Ding, Ye Kyaw Thu, Masao Utiyama, and Eiichiro Sumita. Word segmentation for Burmese (Myanmar). ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(4): 22:1–22:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Zhang:2016:ITP

[42] Tongtao Zhang, Aritra Chowdhury, Nimit Dhulekar, Jinjing Xia, Kevin Knight, Heng Ji, Bülent Yener, and Liming Zhao. From image to translation: Processing the endangered Nyushu script. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15 (4):23:1–23:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Sarigil:2016:SPW

[43] Erdem Sarigil, Oguz Yilmaz, Ismail Sengor Altingovde, Rifat Ozcan, and ÖzgÜr Ulusoy. A "suggested" picture of Web search in Turkish. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(4):24:1–24:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Goswami:2016:CPG

[44] Mukesh M. Goswami and Suman K. Mitra. Classification of printed gujarati characters using low-level stroke features. *ACM Transactions on* 

Asian and Low-Resource Language Information Processing (TALLIP), 15 (4):25:1–25:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Choudhary:2016:FTA

[45] Prakash Choudhary and Neeta Nain. A four-tier annotated Urdu hand-written text image dataset for multidisciplinary research on Urdu script. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(4):26:1–26:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Norimatsu:2016:FCL

[46] Jun-Ya Norimatsu, Makoto Yasuhara, Toru Tanaka, and Mikio Yamamoto. A fast and compact language model implementation using double-array structures. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(4):27:1–27:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Yang:2016:LGF

[47] Haitong Yang and Chengqing Zong. Learning generalized features for semantic role labeling. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 15(4):28:1–28:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Bhowmik:2016:BHC

[48] Tapan Kumar Bhowmik, Swapan Kumar Parui, Utpal Roy, and Lambert Schomaker. Bangla handwritten character segmentation using structural features: a supervised and bootstrapping approach. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 15(4):29:1–29:??, June 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Singh:2016:OHG

[49] Sukhdeep Singh, Anuj Sharma, and Indu Chhabra. Online handwritten Gurmukhi strokes dataset based on minimal set of words. *ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP)*, 16(1):1:1–1:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# El-Fiqi:2016:PCC

[50] Heba El-Fiqi, Eleni Petraki, and Hussein A. Abbass. Pairwise comparative classification for translator stylometric analysis. *ACM Transactions on* 

Asian and Low-Resource Language Information Processing (TALLIP), 16 (1):2:1–2:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Qiao:2016:IUD

[51] Xiuming Qiao, Hailong Cao, and Tiejun Zhao. Improving unsupervised dependency parsing with knowledge from query logs. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(1):3:1–3:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Passban:2016:BNP

[52] Peyman Passban, Qun Liu, and Andy Way. Boosting neural POS tagger for Farsi using morphological information. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16 (1):4:1-4:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Liu:2016:SBM

[53] Liangliang Liu and Cungen Cao. A seed-based method for generating Chinese confusion sets. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(1):5:1–5:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Li:2016:ISP

[54] Junhui Li, Muhua Zhu, Wei Lu, and Guodong Zhou. Improving semantic parsing with enriched synchronous context-free grammars in statistical machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(1):6:1–6:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Krishnamurthi:2016:UDS

[55] Karthik Krishnamurthi, Vijayapal Reddy Panuganti, and Vishnu Vardhan Bulusu. Understanding document semantics from summaries: a case study on Hindi texts. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(1):7:1–7:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Tursun:2016:STT

[56] Eziz Tursun, Debasis Ganguly, Turghun Osman, Ya-Ting Yang, Ghalip Abdukerim, Jun-Lin Zhou, and Qun Liu. A semisupervised tag-transition-based Markovian model for Uyghur morphology analysis. *ACM Transac*-

tions on Asian and Low-Resource Language Information Processing (TAL-LIP), 16(2):8:1–8:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Nguyen:2016:ACN

[57] Long H. B. Nguyen, Dien Dinh, and Phuoc Tran. An approach to construct a named entity annotated English-Vietnamese bilingual corpus. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(2):9:1–9:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Chou:2016:BWN

[58] Chien-Lung Chou, Chia-Hui Chang, and Ya-Yun Huang. Boosted Web named entity recognition via tri-training. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(2): 10:1–10:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Sadek:2016:DBA

[59] Jawad Sadek and Farid Meziane. A discourse-based approach for Arabic question answering. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(2):11:1–11:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Tran:2016:WRS

[60] Phuoc Tran, Dien Dinh, and Long H. B. Nguyen. Word re-segmentation in Chinese-Vietnamese machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(2): 12:1-12:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Li:2016:MSC

[61] Peifeng Li, Guodong Zhou, and Qiaoming Zhu. Minimally supervised Chinese event extraction from multiple views. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16 (2):13:1–13:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Atreya:2016:QER

[62] Arjun Atreya V, Ashish Kankaria, Pushpak Bhattacharyya, and Ganesh Ramakrishnan. Query expansion in resource-scarce languages: a multilingual framework utilizing document structure. ACM Transactions on

Asian and Low-Resource Language Information Processing (TALLIP), 16 (2):14:1–14:??, December 2016. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Finch:2017:IBL

[63] Andrew Finch, Taisuke Harada, Kumiko Tanaka-Ishii, and Eiichiro Sumita. Inducing a bilingual lexicon from short parallel multiword sequences. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):15:1–15:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Wang:2017:CSC

[64] Shaonan Wang and Chengqing Zong. Comparison study on critical components in composition model for phrase representation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):16:1–16:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Bhat:2017:ITB

[65] Riyaz Ahmad Bhat, Irshad Ahmad Bhat, and Dipti Misra Sharma. Improving transition-based dependency parsing of Hindi and Urdu by modeling syntactically relevant phenomena. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):17:1-17:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Das:2017:NER

[66] Arjun Das, Debasis Ganguly, and Utpal Garain. Named entity recognition with word embeddings and Wikipedia categories for a low-resource language. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):18:1–18:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Li:2017:IDR

[67] Haoran Li, Jiajun Zhang, and Chengqing Zong. Implicit discourse relation recognition for English and Chinese with multiview modeling and effective representation learning. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):19:1–19:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Tholpadi:2017:CBT

[68] Goutham Tholpadi, Chiranjib Bhattacharyya, and Shirish Shevade. Corpus-based translation induction in Indian languages using auxiliary language corpora from Wikipedia. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):20:1–20:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Zhao:2017:HMC

[69] Hai Zhao, Deng Cai, Yang Xin, Yuzhu Wang, and Zhongye Jia. A hybrid model for Chinese spelling check. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(3):21:1-21:??, March 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Wali:2017:ECL

[70] Wafa Wali, Bilel Gargouri, and Adelmajid Ben Hamadou. Evaluating the content of LMF standardized dictionaries: a practical experiment on Arabic language. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(4):22:1–22:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Baly:2017:STM

[71] Ramy Baly, Hazem Hajj, Nizar Habash, Khaled Bashir Shaban, and Wassim El-Hajj. A sentiment Treebank and morphologically enriched recursive deep models for effective sentiment analysis in Arabic. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 16(4):23:1–23:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Punchimudiyanse:2017:AFW

[72] Malinda Punchimudiyanse and Ravinda Gayan Narendra Meegama. Animation of fingerspelled words and number signs of the Sinhala sign language. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(4):24:1–24:??, September 2017. CODEN???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Al-Sallab:2017:ARD

[73] Ahmad Al-Sallab, Ramy Baly, Hazem Hajj, Khaled Bashir Shaban, Wassim El-Hajj, and Gilbert Badaro. AROMA: a recursive deep learning model for opinion mining in Arabic as a low resource language. *ACM* 

Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(4):25:1–25:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Kong:2017:CSE

[74] Fang Kong and Guodong Zhou. A CDT-styled end-to-end Chinese discourse parser. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(4):26:1–26:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Liu:2017:PAL

[75] Shih-Hung Liu, Kuan-Yu Chen, Yu-Lun Hsieh, Berlin Chen, Hsin-Min Wang, Hsu-Chun Yen, and Wen-Lian Hsu. A position-aware language modeling framework for extractive broadcast news speech summarization. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(4):27:1–27:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Phani:2017:SLA

[76] Shanta Phani, Shibamouli Lahiri, and Arindam Biswas. A supervised learning approach for authorship attribution of Bengali literary texts. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 16(4):28:1–28:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Passban: 2017: TLR

[77] Peyman Passban, Qun Liu, and Andy Way. Translating low-resource languages by vocabulary adaptation from close counterparts. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 16(4):29:1–29:??, September 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

S:2017:RMI

[78] Sreelekha S. and Pushpak Bhattacharyya. Role of morphology injection in SMT: a case study from Indian language perspective. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(1):1:1–1:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Malik:2017:UNE

[79] Muhammad Kamran Malik. Urdu named entity recognition and classification system using artificial neural network. ACM Transactions on

Asian and Low-Resource Language Information Processing (TALLIP), 17 (1):2:1–2:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Kim:2017:PEN

[80] Hyun Kim, Hun-Young Jung, Hongseok Kwon, Jong-Hyeok Lee, and Seung-Hoon Na. Predictor–estimator: Neural quality estimation based on target word prediction for machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(1):3:1–3:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Almeman:2017:ABV

[81] Khalid Almeman. Automatically building VoIP speech parallel corpora for Arabic dialects. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(1):4:1–4:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Tran:2017:LRB

[82] Phuoc Tran, Dien Dinh, Tan Le, and Long H. B. Nguyen. Linguistic-relationships-based approach for improving word alignment. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(1):5:1–5:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Cheng:2017:ECC

[83] Xiyao Cheng, Ying Chen, Bixiao Cheng, Shoushan Li, and Guodong Zhou. An emotion cause corpus for Chinese microblogs with multiple-user structures. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(1):6:1–6:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Sarma:2017:DAS

[84] Himangshu Sarma, Navanath Saharia, and Utpal Sharma. Development and analysis of speech recognition systems for Assamese language using HTK. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(1):7:1–7:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Bhattacharya:2017:COB

[85] Nilanjana Bhattacharya, Umapada Pal, and Partha Pratim Roy. Cleaning of online Bangla free-form handwritten text. ACM Transactions on

Asian and Low-Resource Language Information Processing (TALLIP), 17 (1):8:1–8:??, November 2017. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Nasution:2018:GCA

[86] Arbi Haza Nasution, Yohei Murakami, and Toru Ishida. A generalized constraint approach to bilingual dictionary induction for low-resource language families. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(2):9:1–9:??, February 2018. CO-DEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Onyenwe:2018:BLR

[87] Ikechukwu E. Onyenwe, Mark Hepple, Uchechukwu Chinedu, and Ignatius Ezeani. A basic language resource kit implementation for the Igbo NLP project. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(2):10:1–10:??, February 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Jia:2018:IDP

[88] Yanyan Jia, Yansong Feng, Yuan Ye, Chao Lv, Chongde Shi, and Dongyan Zhao. Improved discourse parsing with two-step neural transition-based model. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(2):11:1–11:??, February 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Naili:2018:CSS

[89] Marwa Naili, Anja Habacha Chaibi, and Henda Hajjami Ben Ghezala. The contribution of stemming and semantics in Arabic topic segmentation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(2):12:1-12:??, February 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Fujita:2018:EPL

[90] Atsushi Fujita and Pierre Isabelle. Expanding paraphrase lexicons by exploiting generalities. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(2):13:1–13:??, February 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Wang:2018:EEW

[91] Shaonan Wang, Jiajun Zhang, and Chengqing Zong. Empirical exploring word-character relationship for Chinese sentence representation. ACM

Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(3):14:1–14:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Jia:2018:COR

[92] Shengbin Jia, Shijia E., Maozhen Li, and Yang Xiang. Chinese open relation extraction and knowledge base establishment. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(3):15:1–15:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Marie:2018:PTI

[93] Benjamin Marie and Atsushi Fujita. Phrase table induction using monolingual data for low-resource statistical machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP)*, 17(3):16:1–16:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Salami:2018:ISS

[94] Shahram Salami and Mehrnoush Shamsfard. Integrating shallow syntactic labels in the phrase-boundary translation model. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17 (3):17:1-17:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Sherkawi:2018:ASA

[95] Lina Sherkawi, Nada Ghneim, and Oumayma Al Dakkak. Arabic speech act recognition techniques. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(3):18:1–18:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Jung:2018:EEK

[96] Sangkeun Jung, Changki Lee, and Hyunsun Hwang. End-to-end Korean part-of-speech tagging using copying mechanism. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17 (3):19:1–19:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Sen:2018:AST

[97] Shibaprasad Sen, Ankan Bhattacharyya, Pawan Kumar Singh, Ram Sarkar, Kaushik Roy, and David Doermann. Application of structural

and topological features to recognize online handwritten Bangla characters. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(3):20:1–20:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

She:2018:LHD

[98] Xiaohan She, Ping Jian, Pengcheng Zhang, and Heyan Huang. Leveraging hierarchical deep semantics to classify implicit discourse relations via a mutual learning method. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(3):21:1–21:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Mohamed:2018:MSP

[99] Emad Mohamed. Morphological segmentation and part-of-speech tagging for the Arabic heritage. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(3):22:1–22:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Huang:2018:IPK

[100] Degen Huang, Jiahuan Pei, Cong Zhang, Kaiyu Huang, and Jianjun Ma. Incorporating prior knowledge into word embedding for Chinese word similarity measurement. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(3):23:1–23:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Ehsani:2018:CWT

[101] Razieh Ehsani, Ercan Solak, and Olcay Taner Yildiz. Constructing a WordNet for Turkish using manual and automatic annotation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(3):24:1–24:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Huang:2018:LRR

[102] Jizhou Huang, Shiqiang Ding, Haifeng Wang, and Ting Liu. Learning to recommend related entities with serendipity for Web search users. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(3):25:1–25:??, May 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Basiri:2018:WII

[103] Mohammad Ehsan Basiri and Arman Kabiri. Words are important: Improving sentiment analysis in the Persian language by lexicon refining.

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(4):26:1–26:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Suryani:2018:RBS

[104] Arie Ardiyanti Suryani, Dwi Hendratmo Widyantoro, Ayu Purwarianti, and Yayat Sudaryat. The rule-based sundanese stemmer. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 17(4):27:1–27:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# He:2018:DPS

[105] Ruifang He, Yaru Wang, Dawei Song, Peng Zhang, Yuan Jia, and Aijun Li. A dependency parser for spontaneous Chinese spoken language. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(4):28:1–28:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# **Bai:2018:IVS**

[106] Xuefeng Bai, Hailong Cao, and Tiejun Zhao. Improving vector space word representations via kernel canonical correlation analysis. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 17(4):29:1–29:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

# Park:2018:NCI

[107] Taekeun Park and Seung-Hoon Kim. Novel character identification utilizing semantic relation with animate nouns in Korean. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(4):30:1–30:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

#### Wang:2018:GBB

[108] Rui Wang, Hai Zhao, Sabine Ploux, Bao-Liang Lu, Masao Utiyama, and Eiichiro Sumita. Graph-based bilingual word embedding for statistical machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(4):31:1-31:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Hamdi:2018:CCS

[109] Ali Hamdi, Khaled Shaban, and Anazida Zainal. CLASENTI: a class-specific sentiment analysis framework. ACM Transactions on Asian and

Low-Resource Language Information Processing (TALLIP), 17(4):32:1–32:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Wang:2018:DSN

[110] Limin Wang, Shoushan Li, Qian Yan, and Guodong Zhou. Domain-specific named entity recognition with document-level optimization. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17(4):33:1–33:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Komiya:2018:CMA

[111] Kanako Komiya, Masaya Suzuki, Tomoya Iwakura, Minoru Sasaki, and Hiroyuki Shinnou. Comparison of methods to annotate named entity corpora. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(4):34:1–34:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

Zhou:2018:WSP

[112] Deyu Zhou, Zhikai Zhang, Min-Ling Zhang, and Yulan He. Weakly supervised POS tagging without disambiguation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 17 (4):35:1–35:??, August 2018. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic).

### Bhattacharya:2019:UCW

[113] Paheli Bhattacharya, Pawan Goyal, and Sudeshna Sarkar. Using communities of words derived from multilingual word vectors for cross-language information retrieval in Indian languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1): 1:1–1:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3208358.

Li:2019:OAE

[114] Maoxi Li and Mingwen Wang. Optimizing automatic evaluation of machine translation with the ListMLE approach. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):2:1-2:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3226045.

# Su:2019:RSA

[115] Ming-Hsiang Su, Chung-Hsien Wu, Kun-Yi Huang, and Wu-Hsuan Lin. Response selection and automatic message-response expansion in retrieval-based QA systems using semantic dependency pair model. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):3:1–3:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3229184.

### Huang:2019:IMH

[116] Guoping Huang, Jiajun Zhang, Yu Zhou, and Chengqing Zong. Input method for human translators: a novel approach to integrate machine translation effectively and imperceptibly. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1): 4:1-4:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3230638.

### Altakrori:2019:AAA

[117] Malik H. Altakrori, Farkhund Iqbal, Benjamin C. M. Fung, Steven H. H. Ding, and Abdallah Tubaishat. Arabic authorship attribution: an extensive study on Twitter posts. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):5:1-5:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3236391.

### Zhang:2019:WSB

[118] Shaoning Zhang, Cunli Mao, Zhengtao Yu, Hongbin Wang, Zhongwei Li, and Jiafu Zhang. Word segmentation for Burmese based on dual-layer CRFs. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):6:1-6:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3232537.

### Li:2019:IML

[119] Junjie Li, Haoran Li, Xiaomian Kang, Haitong Yang, and Chengqing Zong. Incorporating multi-level user preference into document-level sentiment classification. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):7:1–7:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3234512.

# Jain:2019:UES

[120] Amita Jain, Minni Jain, Goonjan Jain, and Devendra K. Tayal. "UT-TAM": an efficient spelling correction system for Hindi language based on supervised learning. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(1):8:1-8:??, January 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3264620.

# Murthy:2019:INT

[121] Rudra Murthy, Mitesh M. Khapra, and Pushpak Bhattacharyya. Improving NER tagging performance in low-resource languages via multilingual learning. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):9:1–9:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3238797.

# Jarrar:2019:DBM

[122] Mustafa Jarrar, Fadi Zaraket, Rami Asia, and Hamzeh Amayreh. Diacritic-based matching of Arabic words. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2): 10:1-10:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3242177.

### Bhattacharya:2019:SSW

[123] Nilanjana Bhattacharya, Partha Pratim Roy, and Umapada Pal. Substroke-wise relative feature for online Indic handwriting recognition. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):11:1-11:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3264735.

### Mrinalini:2019:PBP

[124] K. Mrinalini, T. Nagarajan, and P. Vijayalakshmi. Pause-based phrase extraction and effective OOV handling for low-resource machine translation systems. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):12:1-12:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3265751.

### Le:2019:LRM

[125] Ngoc Tan Le, Fatiha Sadat, Lucie Menard, and Dien Dinh. Low-resource machine transliteration using recurrent neural networks. *ACM Transac*-

tions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(2):13:1-13:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3265752.

# Na:2019:TBK

[126] Seung hoon Na, Jianri Li, Jong hoon Shin, and Kangil Kim. Transition-based Korean dependency parsing using hybrid word representations of syllables and morphemes with LSTMs. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):14:1-14:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3241745.

# Akhtar:2019:IWE

[127] Md Shad Akhtar, Palaash Sawant, Sukanta Sen, Asif Ekbal, and Pushpak Bhattacharyya. Improving word embedding coverage in less-resourced languages through multi-linguality and cross-linguality: a case study with aspect-based sentiment analysis. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):15:1–15:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3273931.

### Nakamura:2019:WBR

[128] Tatsuya Nakamura, Masumi Shirakawa, Takahiro Hara, and Shojiro Nishio. Wikipedia-based relatedness measurements for multilingual short text clustering. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):16:1–16:??, February 2019. CO-DEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3276473.

#### Ding:2019:NFF

[129] Chenchen Ding, Masao Utiyama, and Eiichiro Sumita. NOVA: a feasible and flexible annotation system for joint tokenization and part-of-speech tagging. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):17:1-17:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3276773.

## Ahmadi:2019:RBK

[130] Sina Ahmadi. A rule-based Kurdish text transliteration system. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):18:1–18:??, February 2019. CODEN ???? ISSN

2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3278623.

### Kamila:2019:THL

[131] Sabyasachi Kamila, Mohammad Hasanuzzaman, Asif Ekbal, and Pushpak Bhattacharyya. Tempo-HindiWordNet: a lexical knowledge-base for temporal information processing. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(2):19:1–19:??, February 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3277504.

# Alnawas:2019:SAI

[132] Anwar Alnawas and Nursal Arici. Sentiment analysis of Iraqi Arabic dialect on Facebook based on distributed representations of documents. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):20:1-20:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ ft\_gateway.cfm?id=3278605.

### Singh:2019:OHG

[133] Sukhdeep Singh and Anuj Sharma. Online handwritten Gurmukhi words recognition: an inclusive study. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):21:1-21:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3282441.

### Yucesoy:2019:COW

[134] Veysel Yücesoy and Aykut Koç. Co-occurrence weight selection in generation of word embeddings for low resource languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(3):22:1-22:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3282443.

#### Bounhas:2019:UCA

[135] Ibrahim Bounhas. On the usage of a classical Arabic corpus as a language resource: Related research and key challenges. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):23:1-23:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3277591.

# Jung:2019:MPN

[136] Sangkeun Jung, Cheon-Eum Park, and Changki Lee. Multitask pointer network for Korean dependency parsing. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):24:1–24:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3282442.

# Bolucu:2019:UJP

[137] Necva Bölücü and Burcu Can. Unsupervised joint PoS tagging and stemming for agglutinative languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):25:1-25:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3292398.

# Kang:2019:SDR

[138] Xiaomian Kang, Chengqing Zong, and Nianwen Xue. A survey of discourse representations for Chinese discourse annotation. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(3):26:1-26:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3293442.

### Badaro:2019:SOM

[139] Gilbert Badaro, Ramy Baly, Hazem Hajj, Wassim El-Hajj, Khaled Bashir Shaban, Nizar Habash, Ahmad Al-Sallab, and Ali Hamdi. A survey of opinion mining in Arabic: a comprehensive system perspective covering challenges and advances in tools, resources, models, applications, and visualizations. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):27:1–27:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3295662.

### Masmoudi:2019:ADR

[140] Abir Masmoudi, Salima Mdhaffar, Rahma Sellami, and Lamia Hadrich Belguith. Automatic diacritics restoration for Tunisian dialect. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):28:1-28:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway. cfm?id=3297278.

#### Rudra:2019:IAD

[141] Koustav Rudra, Ashish Sharma, Kalika Bali, Monojit Choudhury, and Niloy Ganguly. Identifying and analyzing different aspects of English—

Hindi code-switching in Twitter. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):29:1-29:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314935.

# Verma:2019:CAH

[142] Pradeepika Verma, Sukomal Pal, and Hari Om. A comparative analysis on Hindi and English extractive text summarization. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(3):30:1-30:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3308754.

# Wei:2019:ROD

[143] Bingzhen Wei, Xuancheng Ren, Yi Zhang, Xiaoyan Cai, Qi Su, and Xu Sun. Regularizing output distribution of abstractive Chinese social media text summarization for improved semantic consistency. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(3):31:1-31:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314934.

### Trieu:2019:LAR

[144] Hai-Long Trieu, Duc-Vu Tran, Ashwin Ittoo, and Le-Minh Nguyen. Leveraging additional resources for improving statistical machine translation on Asian low-resource languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):32:1–32:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314936.

### Dehghan:2019:CDS

[145] Mohammad Hossein Dehghan and Heshaam Faili. Converting dependency structure into Persian phrase structure. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(3):33:1–33:??, July 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314937.

### Awais:2019:RDI

[146] Muhammad Awais and Muhammad Shoaib. Role of discourse information in Urdu sentiment classification: a rule-based method and machine-learning technique. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):34:1–34:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3300050.

### Nongmeikapam:2019:HMM

[147] Kishorjit Nongmeikapam, Kanan Wahengbam, Oinam Nickson Meetei, and Themrichon Tuithung. Handwritten Manipuri Meetei-Mayek classification using convolutional neural network. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 18 (4):35:1-35:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3309497.

Gao:2019:SBC

[148] Shengxiang Gao, Jihao Huang, Mingya Xue, Zhengtao Yu, Zhuo Wang, and Yang Zhang. Syntax-based Chinese-Vietnamese tree-to-tree statistical machine translation with bilingual features. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18 (4):36:1-36:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314938.

# Sun:2019:NSP

[149] Ruiyong Sun, Yijia Zhao, Qi Zhang, Keyu Ding, Shijin Wang, and Cui Wei. A neural semantic parser for math problems incorporating multi-sentence information. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):37:1-37:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314939.

# Maimaiti:2019:MRT

[150] Mieradilijiang Maimaiti, Yang Liu, Huanbo Luan, and Maosong Sun. Multi-round transfer learning for low-resource NMT using multiple high-resource languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):38:1-38:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314945.

### Ihasz:2019:SFS

[151] Peter Lajos Ihasz, Mate Kovacs, Ian Piumarta, and Victor V. Kryssanov. A supplementary feature set for sentiment analysis in Japanese dialogues. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):39:1–39:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3310283.

### Saeed:2019:SAC

[152] Ali Saeed, Rao Muhammad Adeel Nawab, Mark Stevenson, and Paul Rayson. A sense annotated corpus for all-words Urdu word sense disambiguation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):40:1–40:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314940.

# Dahou:2019:MCE

[153] Abdelghani Dahou, Shengwu Xiong, Junwei Zhou, and Mohamed Abd Elaziz. Multi-channel embedding convolutional neural network model for Arabic sentiment classification. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):41:1-41:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314941.

# Onyenwe:2019:TEI

[154] Ikechukwu E. Onyenwe, Mark Hepple, Uchechukwu Chinedu, and Ignatius Ezeani. Toward an effective Igbo part-of-speech tagger. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):42:1-42:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314942.

### Costa-Jussa:2019:CCN

[155] Marta R. Costa-Jussà, Noé Casas, Carlos Escolano, and José A. R. Fonollosa. Chinese-Catalan: a neural machine translation approach based on pivoting and attention mechanisms. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):43:1-43:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3312575.

### Yu:2019:MTE

[156] Hui Yu, Weizhi Xu, Shouxun Lin, and Qun Liu. Machine translation evaluation metric based on dependency parsing model. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):44:1-44:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id= 3312573.

#### Liu:2019:EBC

[157] Yang Liu, Shaonan Wang, Jiajun Zhang, and Chengqing Zong. Experience-based causality learning for intelligent agents. ACM Transac-

tions on Asian and Low-Resource Language Information Processing (TAL-LIP), 18(4):45:1-45:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3314943.

Yin:2019:PTE

[158] Yongjing Yin, Jinsong Su, Huating Wen, Jiali Zeng, Yang Liu, and Yidong Chen. POS tag-enhanced coarse-to-fine attention for neural machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):46:1–46:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3321124.

Yang:2019:MEA

[159] Jun Yang, Runqi Yang, Hengyang Lu, Chongjun Wang, and Junyuan Xie. Multi-entity aspect-based sentiment analysis with context, entity, aspect memory and dependency information. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18(4):47:1–47:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3321125.

Kim:2019:MTS

[160] Hyun Kim, Jong-Hyeok Lee, and Seung-Hoon Na. Multi-task stack propagation for neural quality estimation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 18(4):48:1–48:??, August 2019. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/ft\_gateway.cfm?id=3321127.

Wang:2020:GCL

[161] Hongmin Wang, Jie Yang, and Yue Zhang. From Genesis to Creole language: Transfer learning for Singlish universal dependencies parsing and POS tagging. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–29, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3321128.

Kong:2020:CZP

[162] Fang Kong, Min Zhang, and Guodong Zhou. Chinese zero pronoun resolution: a chain-to-chain approach. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–21, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3321129.

Yin:2020:CZP

[163] Qingyu Yin, Weinan Zhang, Yu Zhang, and Ting Liu. Chinese zero pronoun resolution: a collaborative filtering-based approach. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 19(1):1-20, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3325884.

Das:2020:TCT

[164] Ayan Das and Sudeshna Sarkar. Transform, combine, and transfer: Delexicalized transfer parser for low-resource languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19 (1):1–30, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3325886.

**Ding:2020:TBM** 

[165] Chenchen Ding, Hnin Thu Zar Aye, Win Pa Pa, Khin Thandar Nwet, Khin Mar Soe, Masao Utiyama, and Eiichiro Sumita. Towards Burmese (Myanmar) morphological analysis: Syllable-based tokenization and part-of-speech tagging. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–34, January 2020. CO-DEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3325885.

Liu:2020:AMC

[166] Dayiheng Liu, Kexin Yang, Qian Qu, and Jiancheng Lv. Ancient-modern Chinese translation with a new large training dataset. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 19(1):1-13, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/ 3325887.

Wang:2020:CSP

[167] Wei Wang, Degen Huang, and Jingxiang Cao. Chinese syntax parsing based on sliding match of semantic string. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–14, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3329707.

Kanwal:2020:UNE

[168] Safia Kanwal, Kamran Malik, Khurram Shahzad, Faisal Aslam, and Zubair Nawaz. Urdu named entity recognition: Corpus generation

and deep learning applications. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–13, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3329710.

Liu:2020:DCW

[169] Yijia Liu, Wanxiang Che, Yuxuan Wang, Bo Zheng, Bing Qin, and Ting Liu. Deep contextualized word embeddings for universal dependency parsing. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–17, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3326497.

# ${\bf Mehmood: 2020: SAR}$

[170] Khawar Mehmood, Daryl Essam, Kamran Shafi, and Muhammad Kamran Malik. Sentiment analysis for a resource poor language —Roman Urdu. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–15, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3329709.

# Bakhshaei:2020:MGM

[171] Somayeh Bakhshaei, Reza Safabakhsh, and Shahram Khadivi. Matching graph, a method for extracting parallel information from comparable corpora. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1-29, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3329713.

### Liu:2020:FTV

[172] Dayiheng Liu, Yang Xue, Feng He, Yuanyuan Chen, and Jiancheng Lv. μ-forcing: Training variational recurrent autoencoders for text generation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–17, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3341110.

### Srivastava:2020:AMA

[173] Jyoti Srivastava, Sudip Sanyal, and Ashish Kumar Srivastava. An automatic and a machine-assisted method to clean bilingual corpus. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(1):1–19, January 2020. CODEN ???? ISSN 2375-4699

(print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3342351.

### Prakash:2020:ISP

[174] Jeena J. Prakash, Golda Brunet Rajan, and Hema A. Murthy. Importance of signal processing cues in transcription correction for low-resource Indian languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–26, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm. org/doi/abs/10.1145/3342352.

### Han:2020:EMW

[175] Dong Han, Junhui Li, Yachao Li, Min Zhang, and Guodong Zhou. Explicitly modeling word translations in neural machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(1):1–17, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3342353.

### Chakrabarty:2020:NNM

[176] Abhisek Chakrabarty, Akshay Chaturvedi, and Utpal Garain. NeuMorph: Neural morphological tagging for low-resource languages-an experimental study for Indic languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(1):1–19, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3342354.

### Ji:2020:ATU

[177] Yatu Ji, Hongxu Hou, Junjie Chen, and Nier Wu. Adversarial training for unknown word problems in neural machine translation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(1):1–12, January 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3342482.

# Zhu:2020:OSK

[178] Qingfu Zhu, Weinan Zhang, Lei Cui, and Ting Liu. Order-sensitive keywords based response generation in open-domain conversational systems. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):18:1–18:18, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3343258.

# Zhou:2020:NCG

[179] Guangyou Zhou, Yizhen Fang, Yehong Peng, and Jiaheng Lu. Neural conversation generation with auxiliary emotional supervised models. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(2):19:1–19:17, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3344788.

# Zhu:2020:EDC

[180] Wenhao Zhu, Xin Jin, Shuang Liu, Zhiguo Lu, Wu Zhang, Ke Yan, and Baogang Wei. Enhanced double-carrier word embedding via phonetics and writing. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):20:1–20:18, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3344920.

# Dehkharghani:2020:SPP

[181] Rahim Dehkharghani. SentiFars: a Persian polarity lexicon for sentiment analysis. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):21:1–21:12, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3345627.

# Abdulhameed:2020:WVT

[182] Tiba Zaki Abdulhameed, Imed Zitouni, and Ikhlas Abdel-Qader. Wasf-Vec: Topology-based word embedding for modern standard Arabic and Iraqi dialect ontology. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):22:1–22:27, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3345517.

### Xu:2020:EPS

[183] Ge Xu, Xiaoyan Yang, Yuanzheng Cai, Zhiqiang Ruan, Tao Wang, and Xiangwen Liao. Extracting polarity shifting patterns from any corpus based on natural annotation. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(2):23:1–23:16, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3345518.

### Imankulova:2020:FPP

[184] Aizhan Imankulova, Takayuki Sato, and Mamoru Komachi. Filtered pseudo-parallel corpus improves low-resource neural machine translation.

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):24:1–24:16, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3341726.

Gupta:2020:DNN

[185] Deepak Gupta, Asif Ekbal, and Pushpak Bhattacharyya. A deep neural network framework for English Hindi question answering. ACM Transactions on Asian and Low-Resource Language Information Processing (TAL-LIP), 19(2):25:1-25:22, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3359988.

Yu:2020:LWT

[186] Hongfei Yu, Xiaoqing Zhou, Xiangyu Duan, and Min Zhang. Layer-wise de-training and re-training for ConvS2S machine translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):26:1–26:15, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3358414.

Somsap:2020:IDW

[187] Sittichai Somsap and Pusadee Seresangtakul. Isarn Dharma word segmentation using a statistical approach with named entity recognition. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):27:1–27:16, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3359990.

Abbas:2020:PIR

[188] Muhammad Raihan Abbas and Dr. Khadim Hussain Asif. Punjabi to ISO 15919 and Roman transliteration with phonetic rectification. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):28:1–28:20, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3359991.

Beseiso:2020:SAM

[189] Majdi Beseiso and Haytham Elmousalami. Subword attentive model for Arabic sentiment analysis: a deep learning approach. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):29:1–29:17, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3360016.

### Harikrishna:2020:CSC

[190] D. M. Harikrishna and K. Sreenivasa Rao. Children's story classification in Indian languages using linguistic and keyword-based features. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):30:1–30:22, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/ abs/10.1145/3342356.

# Jung:2020:WRT

[191] Hun-Young Jung, Jong-Hyeok Lee, Eunju Min, and Seung-Hoon Na. Word reordering for translation into Korean sign language using syntactically-guided classification. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):31:1-31:20, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3357612.

### Masmoudi:2020:TAA

[192] Abir Masmoudi, Mariem Ellouze Khmekhem, Mourad Khrouf, and Lamia Hadrich Belguith. Transliteration of Arabizi into Arabic script for Tunisian dialect. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):32:1-32:21, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3364319.

#### Mukherjee:2020:FST

[193] Subham Mukherjee, Pradeep Kumar, and Partha Pratim Roy. Fusion of spatio-temporal information for Indic word recognition combining online and offline text data. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(2):33:1-33:24, March 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3364533.

### Yu:2020:ELR

[194] Zhiqiang Yu, Zhengtao Yu, Junjun Guo, Yuxin Huang, and Yonghua Wen. Efficient low-resource neural machine translation with reread and feedback mechanism. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):34:1-34:13, February 2020. CO-DEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3365244.

#### Park:2020:NSB

[195] Cheoneum Park, Heejun Song, and Changki Lee.  $S^3$ -NET: SRU-based sentence and self-matching networks for machine reading comprehension.

ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):35:1–35:14, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3365679.

Sarwar:2020:SSF

[196] Raheem Sarwar, Thanasarn Porthaveepong, Attapol Rutherford, Thanawin Rakthanmanon, and Sarana Nutanong. StyloThai:: a scalable framework for stylometric authorship identification of Thai documents. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):36:1–36:15, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3365832.

Kim:2020:UIB

[197] Hyun Kim and Seung-Hoon Na. Uniformly interpolated balancing for robust prediction in translation quality estimation: a case study of English–Korean translation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):37:1-37:27, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3365916.

Zhou:2020:LMU

[198] Xiao Zhou, Zhen-Hua Ling, and Li-Rong Dai. Learning and modeling unit embeddings using deep neural networks for unit-selection-based Mandarin speech synthesis. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):38:1–38:14, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3372244.

Mirzaei:2020:SRL

[199] Azadeh Mirzaei, Fatemeh Sedghi, and Pegah Safari. Semantic role labeling system for Persian language. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):39:1–39:12, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3372246.

Ding:2020:BMT

[200] Chenchen Ding, Sann Su Su Yee, Win Pa Pa, Khin Mar Soe, Masao Utiyama, and Eiichiro Sumita. A Burmese (Myanmar) Treebank: Guideline and analysis. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):40:1–40:13, February

2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3373268.

### Song:2020:KPS

[201] Hyun-Je Song and Seong-Bae Park. Korean part-of-speech tagging based on morpheme generation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):41:1–41:10, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3373608.

# Mi:2020:LIL

[202] Chenggang Mi, Lei Xie, and Yanning Zhang. Loanword identification in low-resource languages with minimal supervision. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19 (3):43:1-43:22, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/ 3374212.

# Li:2020:INM

[203] Yachao Li, Junhui Li, Min Zhang, Yixin Li, and Peng Zou. Improving neural machine translation with linear interpolation of a short-path unit. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):44:1–44:16, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3377851.

### Liu:2020:DUK

[204] Xiao-Yang Liu, Yimeng Zhang, Yukang Liao, and Ling Jiang. Dynamic updating of the knowledge base for a large-scale question answering system. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):45:1–45:13, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3377708.

#### Liu:2020:ELM

[205] Shih-Hung Liu, Kuan-Yu Chen, and Berlin Chen. Enhanced language modeling with proximity and sentence relatedness information for extractive broadcast news summarization. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):46:1–46:19, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3377407.

# Du:2020:CNL

[206] Qianlong Du, Chengqing Zong, and Keh-Yih Su. Conducting natural language inference with word-pair-dependency and local context. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(3):47:1–47:23, February 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3377704.

### Zitouni:2020:ENE

[207] Imed Zitouni. Editorial from the new Editor-in-Chief: the era of natural language processing innovations on Asian and low-resource languages. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):48e:1–48e:2, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3397501.

# Yang:2020:OEQ

[208] Jingxuan Yang, Haotian Cui, Si Li, Sheng Gao, Jun Guo, and Zhengdong Lu. Outline extraction with question-specific memory cells. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):48:1-48:17, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3377707.

### Zarnoufi:2020:MNB

[209] Randa Zarnoufi, Hamid Jaafar, and Mounia Abik. Machine normalization: Bringing social media text from non-standard to standard form. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(4):49:1–49:30, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3378414.

### Bhattu:2020:ICM

[210] S. Nagesh Bhattu, Satya Krishna Nunna, D. V. L. N. Somayajulu, and Binay Pradhan. Improving code-mixed POS tagging using code-mixed embeddings. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):50:1-50:31, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl. acm.org/doi/abs/10.1145/3380967.

### Ahmad:2020:NER

[211] Muhammad Tayyab Ahmad, Muhammad Kamran Malik, Khurram Shahzad, Faisal Aslam, Asif Iqbal, Zubair Nawaz, and Faisal Bukhari.

Named entity recognition and classification for Punjabi Shahmukhi. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(4):51:1–51:13, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3383306.

Garg:2020:HES

[212] Kanika Garg and D. K. Lobiyal. Hindi EmotionNet: a scalable emotion lexicon for sentiment classification of Hindi text. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):52:1–52:35, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3383330.

# Udomcharoenchaikit:2020:AER

[213] Can Udomcharoenchaikit, Prachya Boonkwan, and Peerapon Vateekul. Adversarial evaluation of robust neural sequential tagging methods for Thai language. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):53:1–53:25, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3383201.

### Sugandhi:2020:SLG

[214] Sugandhi, Parteek Kumar, and Sanmeet Kaur. Sign language generation system based on Indian sign language grammar. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):54:1–54:26, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3384202.

### Sarwar:2020:NLI

[215] Raheem Sarwar, Attapol T. Rutherford, Saeed-Ul Hassan, Thanawin Rakthanmanon, and Sarana Nutanong. Native language identification of fluent and advanced non-native writers. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):55:1–55:19, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3383202.

# Buyuk:2020:CDS

[216] Osman Büyük. Context-dependent sequence-to-sequence Turkish spelling correction. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):56:1-56:16, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3383200.

# Khalil:2020:EAC

[217] Hussein Khalil, Taha Osman, and Mohammed Miltan. Extracting Arabic composite names using genitive principles of Arabic grammar. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):57:1–57:16, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3382187.

# Wang:2020:SCH

[218] Kexin Wang, Yu Zhou, Jiajun Zhang, Shaonan Wang, and Chengqing Zong. Structurally comparative hinge loss for dependency-based neural text representation. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):58:1–58:19, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3387633.

# Liu:2020:JME

[219] Maofu Liu, Yukun Zhang, Wenjie Li, and Donghong Ji. Joint model of entity recognition and relation extraction with self-attention mechanism. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 19(4):59:1–59:19, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3387634.

### Kumar:2020:LGV

[220] H. R. Shiva Kumar and A. G. Ramakrishnan. Lipi Gnani: a versatile OCR for documents in any language printed in Kannada script. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*, 19(4):60:1–60:23, July 2020. CODEN ???? ISSN 2375-4699 (print), 2375-4702 (electronic). URL https://dl.acm.org/doi/abs/10.1145/3387632.