Procesamiento del Lenguaje Natural, Revista nº 60, marzo de 2018, pp. 53-60 recibido 22-11-2017 revisado 16-12-2017 aceptado 20-02-2018

Sequential dialogue act recognition for Arabic argumentative debates

Reconocimiento de acto de diálogo secuencial para debates argumentativos árabes

Samira Ben Dbabis[1], Hatem Ghorbel[2], Lamia Hadrich Belguith[3]

[1,3] ANLP Research Group, MIRACL Laboratory, University of Sfax, Tunisia
[2] University of Applied Science of West Switzerland HE-Arc Ingénierie, Switzerland

[1] samira.benedbabis@fsegs.rnu.tn
[2] hatem.ghorbel@he-arc.ch
[3] l.belguith@fsegs.rnu.tn

Abstract: Dialogue act recognition remains a primordial task that helps user to automatically identify participants' intentions. In this paper, we propose a sequential approach consisting of segmentation followed by annotation process to identify dialogue acts within Arabic politic debates.To perform DA recognition, we used the CARD corpus labeled using the SADA annotation schema. Segmentation and annotation tasks were then carried out using Conditional Random Fields probabilistic models as they prove high performance in segmenting and labeling sequential data. Learning results are notably important for the segmentation task (F-score=97.9%) and relatively reliable within the annotation process (f-score=63.4%) given the complexity of identifying argumentative tags and the presence of disfluencies in spoken conversations.
Keywords: DA recognition, annotation scheme, Arabic debates, CRF classifier.

Resumen: El reconocimiento del acto de diálogo sigue siendo una tarea primordial que ayuda al usuario a identificar automáticamente las intenciones de los participantes. En este documento, proponemos un enfoque secuencial que consiste en la segmentación seguida de un proceso de anotación para identificar actos de diálogo dentro de los debates políticos árabes. Para realizar el reconocimiento DA, utilizamos el corpus CARD etiquetado utilizando el esquema de anotación SADA. Las tareas de segmentación y anotación se llevaron a cabo utilizando modelos probabilísticos de Campos aleatorios condicionales, ya que demuestran un alto rendimiento en la segmentación y el etiquetado de datos secuenciales. Los resultados de aprendizaje son especialmente importantes para la tarea de segmentación (F-score = 97.9%) y relativamente confiables dentro del proceso de anotación (f-score = 63.4%) dada la complejidad de identificar etiquetas argumentativas y la presencia de disfluencias en las conversaciones habladas.
Palabras clave: Reconocimiento DA, esquema de anotación, debates árabes, clasificador CRF.

1 Introduction

Dialogue acts (DA) are considered as the minimal units of linguistic communication that reveal speaker's intention (Grosz and Sidner, 1986). Automatic dialogue act detection is an important clue for various applications like dialogue systems, human conversations understanding, machine translation, topic detection and summarization.

In our work, the aim of dialogue acts recognition is to better understand human conversations based mainly on argumentative tags in order to extract participants' conflicts in terms of opinions' reject or accept and arguments presented to defund their ideas.

ISSN 1135-5948. DOI 10.26342/2018-60-6

To perform this task, we propose in a first step a complete annotation scheme consisting of 40 DAs. In a second step, we reduced the initial scheme to 19 acts as we decided to focus mainly on argumentative tags and merge others for instance social obligation management and turn management categories.

The proposed DAs are automatically identified using machine learning techniques applied on a large corpus collected from politic debates found to have explicit argumentative taxonomy and forms like opinions, arguments, acceptations, rejects, explanations, justifications, etc.

This paper is organized as follows. The first section describes the major DA recognition approaches. In section 2 we detail the implications of DA recognition explored mainly in building argumentative discourse structure. In sections 3 and 4, we present the proposed annotation scheme and the corpus used to perform learning machine experiments. Section 5 details the proposed recognition sequential approach consisting of two main tasks: segmentation followed by annotation of dialogue acts. For each task, we focus on the used learning technique, the experimental data, the adopted features and the evaluation results.

2 Building Argumentative Structure

Dialogue acts play a vital role in the identification of discourse structure. In this context, Grosz and Sidner (1986) claim about task structure influencing dialogue structure. It seems likely that there are structures higher than a single utterance, yet more fine grained than a complete dialogue. Several researchers identify structures within dialogue at levels higher than individual utterances or speaker turns, but below the level of complete discourse description. There has been some significant exploration of the use of sequences of Dialogue Acts, at a number of levels of granularity.

The simplest dialogue sequence model is the use of adjacency pairs (Schegloff et al., 1973) which are functional links between pairs of utterances such as question/answer, opinion request/opinion, etc.

Within the adjacency pairs model, the importance of tracking a deeper structured representation has been recognized in Ezen-Can and Boyer (2015), Swapna and Wiebe (2010) and Galley et al. (2004).

In fact, Ezen-Can and Boyer (2015) investigate sequences of acts to automatically detect the interaction mode between students and teachers (tutor lecture, tutor evaluator, Extra-domain and student). Swapna and Wiebe (2010) use the AMI corpus (Carletta et al., 2005) to detect opinions' categories such argument and sentiment in meetings. Galley et al. (2004) also explored adjacent act chains to extract the agreement and disagreement pairs within meetings of the ICSI corpus (Shriberg et al., 2004).

In our work, the main implications of recognizing dialogue acts are to build argumentative chains consisting of pairs or more than two acts to highlight argumentative interaction between participants. For instance an opinion request asked by the animator is generally followed by an opinion tag which can be rejected or accepted by other participants. The opinion holder can reinforce his point of view by exposing arguments, explanations or justifications.

Thus, dialogue act sequences can help in capturing the essential argumentative information in terms of what topics have been discussed and what alternatives have been proposed and accepted by the participants. They can be also useful in opinion question/answering systems to answer complex real user queries like "who rejected the opinion of X?" which is not evident to reply using traditional information retrieval engines.

3 Annotation scheme

Over the years a number of dialogue act annotation schemas has been developed, such as those of the MapTask studies outlining road mapping task-oriented dialogues (Carletta, 1996) and the Verbmobil project (Alexandersson et al., 1998) focusing on meeting scheduling and travel planning domains. Later, DAMSL (Core, and Allen, 1997) annotation schema was developed for multidimensional dialogue act annotation. As an extension of DAMSL, The DIT++ schema (Bunt, 2009) combines the multidimensional DIT schema, developed earlier (Bunt, 1994) with concepts from these various alternative schemas, and provides precise and mutually consistent definitions for its communicative functions and dimensions.

These annotation schemes have been used to mark-up several dialogue corpora in non Arabic languages. To the best of our knowledge, few works were developed in Arabic language. We mention the taxonomy proposed by Shala et al. (2010) that proposed speech acts taxonomy including the following set of 10 categories dealing with general information requests followed by answers.

Recently, Elmadany et al. (2014) reported a schema for inquiry-answer instant messages in Egyptian dialect such as flights, mobile service operators, and banks; this schema contains 25 DAs based on request and response dimensions.

Given that the main purpose of identifying dialogue acts is to build argumentative discourse structure, we cannot profit from previous annotation schemes and we need to develop a specific-purpose taxonomy based mainly on argumentative acts called SADA: **S**cheme **A**nnotation for **D**ebates in **A**rabic.

The first release of SADA (BenDbabis et al., 2012) is a complete tagset consisting of 40 dialogue acts related to the following categories: *social obligation management*, *turn management*, *Request*, *Argumentative*, *Answer, statement and others.*

In a second step, we reduced the initial tagset to 19 acts (BenDbabis et al., 2015). We merge acts expressing social obligation management into a single dialogue act named *SOM*. We also combine acts expressing Turn Management in one act labeled *TM.* We eliminate acts having very few occurrences in the corpus like *statement*, *propose, hope, wish, invoke, warn* and *order.*

We also eliminate the following tags expressing Appreciation (*app*), disapproval (*disap*), partial accept (*part_acc*) and partial reject (*part_rej*). In fact, we considered appreciation and partial accept acts as acceptation tags while disapproval and partial reject was considered as forms of reject. We add the tag *Thesis* in the argumentative category referring to a new topic or idea introduced by the presenter that can be retained or rejected by the audience.

4 CARD corpus

Corpora annotated for Dialogue Acts play a key role in the validation and evaluation of the proposed annotation taxonomies. In our context of work, our main purpose is to track argumentative information from human conversations. Thus, we collected a set of politic debates from Aljazeera TV broadcasts discussing hot topics (Tunisian and Egyptian revolutions, Syrian war, Tunisian elections, etc); named CARD: **C**orpus of **AR**abic **D**ebates. The choice of this corpus is argued by the important argumentation hold in its content mainly conveyed by exchanging opinions, agreements, disagreements, etc.

The CARD corpus was manually annotated using the ActAAr annotation tool: **Ac**t **A**nnotation in **Ar**abic (BenDbabis et al., 2012) in three steps reaching 50 conversations in the latest release. Basic information of the different versions of the CARD corpus is detailed in Table 1.

	CARD 1.0	**CARD 1.1**	**CARD 1.2**
Total number of conversations	8	22	50
Total number of turns	773	1805	5085
Total number of utterances	2367	6050	14062
Total number of words	37075	101169	260212
Average number of turns/conversation	97	82	102
Average number of utterances/conversation	296	275	281
Average number of words/conversation	4635	4599	5204

Table 1: CARD corpus statistics

5 DA recognition

Dialogue act recognition consists mainly of two subtasks as segmentation and annotation. These two steps may be carried separately; segmentation followed by annotation or simultaneously at one joint step. In our work, we typically assumed that the true segmentation boundaries lead to better annotation results. As a consequence, a degradation of the performance due to imperfect segmentation boundaries is to be expected. Thus, we decided to carry out a sequential approach that separate the two subtasks of dialogue acts recognition framework.

5.1 Segmentation task

The Segmentation task consists of dividing the conversation into turns; each turn is then segmented into meaningful units named utterances. For each utterance, a dialogue act unit is assigned. The problem of identifying utterance boundaries has been addressed with machine learning approaches. Most researchers applied generative models Hidden Markov Models (HMM) experimented by Ivanovic (2005) to find the most likely segment boundaries in online instant messages based services and Naïve Bayes generative classifier (Geertzen et al., 2007) for assistance-seeking Dutch dialogues within the DIAMOND corpus.

Discriminative models have been experimented to perform better than HMMs and maximum entropy approaches for utterance segmentation. The most common discriminative models are Conditional Random Fields (CRF) introduced by Lafferty et al. (2001). It was applied by Silvia et al. (2011) using two corpora namely SWITCHBOARD and LUNA corpus.

Semi-supervised learning approaches were also implemented in the purpose to reduce the amount of labeled data needed to train statistical models.

In this context, Guz et al. (2010) applied self-training and co-training approaches using the ICSI meeting corpus (Schrieberg et al., 2004) of multichannel conversational speech data.

Most of utterance segmentation researches were applied on various languages corpora like English, German and Italian. Few works focus on utterance segmentation of Arabic conversations. We cite the work of Elmadany et al. (2015) who proposed an automatic segmentation utterance approach using SVM classifier for Egyptian instant messages.

In our work, we applied the probabilistic CRF learner to automatically define utterances boundaries. The choice of this model is justified by its efficiency for labeling and segmenting sequential data.

To perform training and test tasks, we used the CARD corpus enhanced in three steps ranging from 8 to 50 conversations. BenDbabis et al. (2016) expose utterance segmentation experiments using CARD 1.1 corpus.

5.1.1 Features selection

Selecting most pertinent features has a great effect on learning machine process mainly on resulting labeled data. In our work, we explored lexical features namely punctuation marks and cue words as important indicators of segment boundaries. We also use morpho-syntactic features as the Part Of Speech (POS) of words. For each word, we take into account a context window of +2/-2; that means we consider dependencies between the current word and the two previous and next words.

As a lexical characteristic, we focus on punctuation as a determinant clue that occurs frequently and the end of an utterance. For example question marks mostly delimit the end of a question.

Question words are also considered as pertinent cue words that express a request or a general question in the beginning of conversation segments.

Lexical cues are frequently used to identify the beginning of a segment. For instance the words “مرحبا”/”welcome”, “ أهلا بكم”/”Hello”, “طيب”/”ok”, “نعم”/“yes”, “أوافق”/”I agree” occur generally at first of utterances.

The POS of each word can also help to recognize utterance delimiters given that utterances often start with prepositions (“في”/ “in”, “من”/”from”), adverbs (“طيب”/”ok”, ” أولا”/”first”) or verbs (”أرى”/“I see”, “أعتقد”/ “I think”).

5.1.2 Results

We experiment the CRF classifier using the different CARD versions. For each release of the corpus, we assess precision, recall and f-measure traditional evaluation metrics. To better evaluate CRF efficiency, learning results were compared to SVM, Naïve Bayes (NB) and Decision trees (J48) classifiers. Comparison results of the used classifiers are shown in Figure 1.

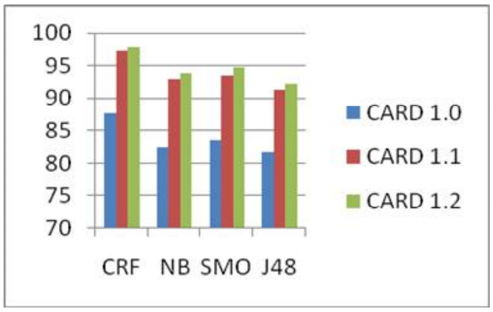

Figure 1: CARD f-measure results

Evaluation results prove the high performance of CRF models in segmenting conversations into meaningful utterances. We obtained a recall rate of 98%, a precision value of 97,8% and an f-measure score of 97,9% using the CARD 1.2 corpus. We also confirm the importance of amount of data in learning machine experiments. Best results are acquired when using larger corpus (CARD 1.2) including 260383 words.

Segmentation errors are mainly due to the fact that punctuation marks can be used inside segments. Mis-segmented utterances can be also explained by the presence of cue words inside utterances which can lead to wrong segment boundaries.

5.2 Annotation task

The annotation task is fundamental in dialogue acts recognition framework. For each segmented utterance, we assign a label expressing the user's intention throughout the conversation. Research has continued to experiment machine learning techniques to automatically identify DAs. Supervised modeling approaches are frequently used including sequential approaches and vector-based models.

Sequential approaches typically formulate dialogue as a Markov chain in which an observation depends on a finite number of preceding observations. HMM-based approaches generate optimal dialogue act sequences using the Viterbi (Stolcke et al., 2000; Bangalore et al., 2008; Ondáš et al., 2016). Research using sequential approaches usually involves combinations of N-grams and Hidden Markov Models.

Vector-based approaches such as maximum entropy (Sridhar et al., 2007) and SVM models (Zhou et al., 2015) frequently take into account lexical, syntactic and structural features. Lexical and syntactic cues are extracted from local utterance context, while structural features involve longer dialogue act sequences in task-oriented domains.

Neuronal networks (Shen et al., 2016) were also investigated to automatically classify dialogue acts. Zhou et al. (2015) applied a combination of heterogeneous deep neural networks with conditional random fields for Chinese corpus.

More interestingly, researchers focused on features enhanced dialogue context (Webb et al., 2005; Hoque et al., 2007; Coria el al., 2007) ;Di Eugenio et al., 2010a; Samei et al., 2014; Ribeiro et al., 2015) that shows a predictive power on Dialogue Act classification. Recently semantic information was explored in the annotation of Czech dialogue corpus (Pavel et al., 2015). Yeh (2016) also involve using semantic dependency graphs with probabilistic context-free grammars (PCFGs).

Most DA annotation classifiers were experimented using several dialogue corpora in different languages such as English, German and Spanish.

However, very few works were developed for Arabic language. Shala et al. (2010) propose speech acts classification model using SVM for the labeling of a tagset of 10 acts. This tagset includes general-purpose actions that can be applied to independent domain corpora.

Elmadany et al. (2015) also experiment SVM model for question-inquiry dialogue acts recognition with a reduced labeling schema of 25 acts for Egyptian spontaneous dialogues and instant messages.

Nevertheless, the proposed annotation works mainly label short utterances expressing requests, questions and answers that are not complex to identify especially with the presence of a predefined list of cue words.

In our work, we implement the CRF model to label utterances segmented in the previous task. The proposed model takes advantage of dependencies between interconnected annotations compared to conventional classification models.

To perform the annotation process, we used the CARD corpus annotated with the SADA annotation schema.

5.2.1 Features selection

DA classification involves linguistic, prosodic and multimodal features. Most of researches explore linguistic features that include lexical, syntactic, semantic and context-based features (Sridhar et al., 2009; Kim et al., 2012). In our context, we choose the most relevant characteristics to our task namely lexical, morpho-syntactic, utterance and structural learning features. We detail below the selected features.

- *Question words:* expressing requests and general questions; for example the word "لماذا"/ " why" indicates a justification request.
- *Cue words :* are most common words frequently used along the conversation. For instance " أهلا بكم" /"welcome", and " شكرا"/"thanks" are used for introducing social obligation management acts.
- *Opinion words:* used when presenting argumentative information like opinions, arguments, acceptations and rejects.
- *Part Of Speech:* grammatical categories of words can reflect the act expressed; for

instance, verbs are frequently used for argumentative tags.

- *Utterance speaker:* the actor of the current utterance.
- *Speaker role:* whether the speaker is the animator of the discussion or just a participant. *Mostly, the animator introduces and ends the discussion and manages the participants' turn taking.*
- *Previous act*: can help to anticipate the next DA label. *For instance, a confirmation request is generally followed by a confirmation.*
- *Previous utterance speaker:* it is important to identify whether the previous utterance has the same actor as the current one.

5.2.2 Results

To train the CARD data, we defined a template that includes unigrams and bigrams of features to focus on the dependencies between features. For each feature, we take into account the two previous and next words (context window=2). We used the CRF++ platform to train and test the CRF model. This tool is an implementation of CRF for labeling sequential data.

Annotation relevance is evaluated using the known metrics as recall, precision and f-measure. All evaluation results shown below were carried out using 10 folds cross validation. To evaluate CRF performance, we compared the obtained results to Naïve Bayes (NB), Decision tree (J48) and SMO classifiers in Figures 2, 3 and 4. For all classifiers, we notice that the increase of the corpus size improves notably the annotation results. For instance, CRF achieves an f-score of 32,4% with CARD 1.0 while the latest release CARD 1.2 reaches an f-measure rate of 63,4% using the same classifier.

Annotation results show that CRF model outperform other classifiers with all releases of CARD corpus with a recall value of 62,2%, a precision rate equal to 64,7% and an f-measure of 63,4%. Thus, CRF results reinforce the high performance of this classifier in labeling sequential data.

Main annotation errors are due classification ambiguities for identifying argumentative tags. There are confusions between arguments, explanations and opinion acts especially when specific lexical cues are absent. Turn management utterances are generally predicted by the enunciation context. So it is difficult to identify these tags that don't obey to the use of general rules or particular cue words. In addition, some lexical cues can have different meaning depending on context. For example, the word "ok" can be used as a form of acceptation or as an acknowledgement act to manage the conversation turn takings.

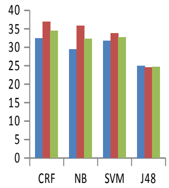

Figure 2: CARD 1.0

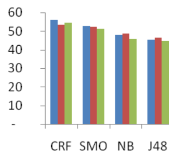

Figure 3: CARD 1.1

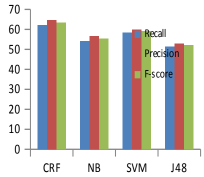

Figure 4: CARD 1.2

6 Conclusion and perspectives

To the best of our knowledge there is no similar work that identifies argumentative dialogue acts within politic debates. In this paper, we proposed a novel sequential dialogue act recognition approach carrying out separately segmentation and annotation tasks. To automatically perform dialogue act identification process, we applied the probabilistic model CRF in both segmentation and labeling subtasks. Results confirm the effectiveness of CRF compared to naïve bayes, SVM and decision trees learning algorithms. Annotation experiments are very encouraging with an average F-score of 63,4%. These results are due to the complexity of labeling argumentative information and difficulties to differentiate between corresponding acts which can need a pragmatic level to enhance the recognition process.

As future work, we intend to integrate context-based and semantic features to improve the annotation results. We also project to investigate the annotated dataset in an extrinsic task such opinion question answering, argumentative discourse structure building and conversations summarization.

7 References

Alexandersson J., B. Buschbeck-Wolf, T. Fujinami, M. Kipp, S. Koch, E. Maier, N. Reithinger, B. Schmitz, and M. Siegel. 1998. Dialogue Acts in VERBMOBIL-2 (second edition). *Vm report 226*, DFKI GmbH, Universities of Berlin, Saarbrcken and Stuttgart.

Bangalore S., G. Di Fabbrizio and A. Stent. 2008. Learning the structure of task-driven human–human dialogs. *IEEE Transactions on Audio, Speech, and Language Processing*. 16(7):1249–1259.

BenDbabis S., H. Ghorbel, L. Belguith and M. Kallel. 2015. Automatic dialogue acts Annotation within Arabic debates. *16th International Conference on Intelligent Text Processing and Computational Linguistics* , April 14-20, Cairo, Egypt.

BenDbabis S., B. Rguii, H. Ghorbel and L. Belguith. 2016. Utterance Segmentation Using Conditional Random Fields. *27th International Business on Information Management Association*, May 3-5, 2016, Milano, Italy.

BenDbabis S., F. Mallek, H. Ghorbel and L. Belguith. 2012. Dialogue Acts Annotation Scheme within Arabic discussions. *Sixteenth Workshop on the Semantics and Pragmatics of Dialogue,* September 19-21, Paris, France.

Bunt H. 2009. The DIT++ taxonomy for functional dialogue markup. *In Proceedings of the AAMAS Workshop*, Budapest, May 12, 2009.

Bunt H. 1994. Context and Dialogue Control. THINK, 3:19-31.

Carletta J., S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, W. Kraaij, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, I. McCowan, W. Post, D. Reidsma, and P. Wellner. 2005. The AMI Meetings Corpus. *In Proceedings of the Measuring Behavior Symposium on "Annotating and measuring Meeting Behavior"*.

Carletta J. C. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22(2): 249-254.

Core M. and J. Allen. 1997. Coding Dialogs with the DAMSL annotation scheme. *In AAAI Fall Symposium on Communicative Action in Humans and Machines*, MIT, Cambridge, MA.

Coria S. R. and L. A. y Pineda. 2007. Prediction of Dialogue Acts on the Basis of Previous Act. *Procesamiento de Lenguaje Natural,* 39: 223 – 230.

Di Eugenio B., Z. Xie and R.o Serafin. 2014. Dialogue act classification, higher order dialogue structure, and instance-based learning. Dialogue and Discourse,1(2):1–24.

Elmadany A., S. M. Abdou and M. Gheith. 2015. Turn Segmentation Into Utterances For Arabic Spontaneous Dialogues And Instant Messages. *International Journal on Natural Language Computing,* 4(2). April 2015.

Elmadany A., S. M. Abdou, and M. Gheith. 2014. Arabic Inquiry-Answer Dialogue Acts Annotation Schema. *IOSR Journal of Engineering*, 4(12-V2):32-36.

Ezen-Can A. and K. E. Boyer. 2015. A tutorial dialogue system for real-time evaluation of unsupervised dialogue act classifiers: exploring system outcomes. *In International Conference on Artificial Intelligence in Education*. Springer International Publishing. Pages 105-114.

Galley M., K. McKeown, J. Hirschberg, and E. Shriberg. 2004. Identifying agreement and disagreement in conversational speech: Use of bayesian networks to model pragmatic dependencies. *In ACL 2004*, Barcelona.

Geertzen J., V. Petukhova and H. Bunt. 2007. A Multidimensional Approach to Utterance Segmentation and Dialogue Act Classification. *In Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*, Antwerp.

Grosz B. J. and C. L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*. 12(3): 175-204.

Guz U., S. Cuendet, D. Hakkani-Tür and G. Tur. 2010. Multi-View Semi-Supervised Learning for Dialog Act Segmentation of Speech. *IEEE Transactions on Audio, Speech and Language Processing*: pages 320-329.

Hoque M. E., M. S. Sorower, M. Yeasin and M. M. Louwerse. 2007. What Speech Tells us about Discourse: The Role of Prosodic and Discourse Features in Dialogue Act Classification. *In IEEE International Joint Conference on Neural Networks (IJCNN)*, Orlando, FL.

Ivanovic E. 2005. Automatic Utterance Segmentation in Instant Messaging Dialogue. *Proceedings of the Australasian Language Technology Workshop*. Pages 241–249, Sydney, Australia.

Lafferty J., A. McCallum and F. Pereira. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. *In Proceedings of the 18th International Conference on Machine Learning*, pages 282-289.

Ondáš S. and J. Juhár. 2016. Towards human-machine dialog in Slovak . *International Conference on Systems, Signals and Image Processing (IWSSIP)*, Bratislava, pages 1-4.

Pavel K., L. Lenc and C. Cerisara. 2015. Semantic Features for Dialogue Act Recognition. *Statistical Language and Speech Processing*. Springer International Publishing, pages 153-163.

Kim S., L. Cavedon and T. Baldwin. 2012. Classifying dialogue acts in multi-party live chats. *In Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*, pages 463–472, Bali, Indonesia, November.

Ribeiro E., R. Ribeiro and D. M. de Matos. 2015. The Influence of Context on Dialogue Act Recognition. arXiv preprint arXiv:1506.00839.

Samei B., H. Li, F. Keshtkar, V. Rus, and A. C. Graesser. 2014. Context-based speech act classification in intelligent tutoring systems. *In International Conference on Intelligent Tutoring Systems*, pages 236-241. Springer International Publishing.

Schegloff E. A. and H. Sacks. 1973. Opening Up Closings. Semiotica, 7:289-327.

Shala, V. Rus and A. C. Graesser. 2010. L. Automated speech act classification in Arabic. *Subjetividad y Procesos Cognitivos*, 14: 284-292.

Shen S. S. and H. Y. Lee. 2016. Neural Attention Models for Sequence Classification: Analysis and Application to Key Term Extraction and Dialogue Act Detection. arXiv preprint arXiv:1604.00077.

Shriberg E., R. Dhillon, S. Bhagat, J. Ang and H. Carvey. 2004. The ICSI meeting recorder dialog act (MRDA) corpus. *In Proceedings of the 5th SIGDIAL Workshop on Discourse and Dialogue*.

Silvia Q., V. Alexei and R. Giuseppe. 2011. Simultaneous dialog act segmentation and classification from human-human spoken conversations. *ICASSP 2011*: pages 5596-5599.

Sridhar V.K.R., S. Bangalore and S.S. Narayanan. 2007. Exploiting acoustic and syntactic features for prosody labeling in a maximum entropy framework. *NAACL-HLT*.

Sridhar V.K.R, S. Bangalore, and S.S. Narayanan. 2009. Combining lexical, syntactic and prosodic cues for improved online dialog act tagging. *Computer Speech and Language*, 23(4): 407-422. Elsevier Ltd.

Stolcke A., K. Ries, N. Coccaro, E. Shriberg, R. Bates, D. Jurafsky, P. Taylor, R. Martin, C. Van Ess-Dykema, and M. Meteer. 2000. Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech. *In Computational Linguistics 2000*. 26(3): 339-373.

Swapna S. and J. Wiebe. 2010. Recognizing Stances in Ideological On-line Debates. *In Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116-124, Los Angeles, CA.

Webb N., M. Hepple and Y. Wilks. 2005. Dialogue act classification based on intra-utterance features. *In Proceedings of the AAAI Workshop on Spoken Language Understanding*, Pittsburgh, PA.

Yeh J.. 2016. Speech Act Identification Using Semantic Dependency Graphs with Probabilistic Context-Free Grammars. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 15(1).

Zhou Y., Q. Hu, J. Liu, and Y. Jia. 2015. Combining heterogeneous deep neural networks with conditional random fields for Chinese dialogue act recognition. *Neurocomputing*, 168: 408-417.