CLiC-it 2017: A Retrospective

Roberto Basili*
Università di Roma, Tor Vergata

Malvina Nissim**
Rijksuniversiteit Groningen

Giorgio Satta†
Università di Padova

The Fourth Italian Conference on Computational Linguistics (CLiC-it 2017) took place in Rome, in December 2017. As in previous editions, it served as the prime forum in Italy for discussing research in computational linguistics and Natural Language Processing. As General Chairs, we offer a retrospective over the conference, highlighting its international flavour and its attention to students and young researchers, with a particular focus on the innovations that were introduced at the 2017 edition.

1. Context

The Fourth Italian Conference on Computational Linguistics (CLiC-it) took place on December 11–13, 2017, with more than 110 registered participants and for the first time in the wonderful city of Rome. The conference was locally organised by the University of Rome "Tor Vergata", and was hosted at the headquarters of the National Research Council of Italy (CNR). The CLiC-it conference series is an initiative of the Italian Association for Computational Linguistics (AILC) and, after four years of activity, it has clearly established itself as the premier national forum for research and development in the fields of Computational Linguistics and Natural Language Processing (CL/NLP), where leading researchers and practitioners from Italian academia and industry meet to share their research results, experiences, and challenges.

These annual meetings have explicitly allowed reports on ongoing research, with the goal of ensuring a wide participation of the community and motivated by an inclusive spirit. Usually, the number of submitted papers is over 60, with over 100 registered participants. A number of submissions are also accepted as poster presentations. To make participation even more attractive, some internationally well-known researchers from abroad are invited for a keynote lecture, and some panel events are also added to the program. Overall, this event structure has made it possible that "all" researchers in computational linguists in Italy meet together once every year.

2. Technical Program

The conference received 72 submissions, against 64 submissions in 2015 and 69 submissions in 2016. The Programme Committee worked very hard to ensure that every paper

* Department of Enterprise Engineering, University of Rome, Tor Vergata, Via del Politecnico, 1, 00133 Roma (Italy). E-mail: basili@info.uniroma2.it
** CLCG, Rijksuniversiteit Groningen, Groningen NL m.nissim@rug.nl
† Department of Information Engineering, University of Padua, Italy. E-mail: satta@dei.unipd.it

Table 1
Areas at the CLiC-it 2017, and number (proportion) of accepted papers per area.

AREA	ACCEPTED	
Cognitive Modeling of Language Processing and Psycholinguistics	3	5.2%
Information Extraction, Information Retrieval and Question Answering	3	5.2%
Language Resources	9	15.5%
Linguistic Issues in Computational Linguistics & Natural Language Processing	7	12.1%
Machine Learning and Language Processing	6	10.3%
Machine Translation and Multilinguality	3	5.2%
Morphology and Syntax Processing	2	3.4%
Natural Language Processing for Digital Humanities	6	10.3%
Natural Language Processing for Web and Social Media	6	10.3%
Pragmatics and Creativity	8	13.8%
Semantics and Knowledge Acquisition	3	5.2%
Spoken Language Processing and Automatic Speech Understanding	2	3.4%

received at least two careful and fair reviews. This process finally led to the acceptance of 21 papers for oral presentation and 37 papers for poster presentation, with a global acceptance rate of 80%, again in line with previous editions (81% in 2015 and 80% in 2016). Regardless of the format of presentation, all accepted papers are allocated 6 pages in the proceedings, available as open access publication.[1]

The conference was organised around 12 thematic areas that are basically the same as those of the 2016 edition of CLiC-it, with the only exception of the area Information Retrieval and Question Answering and the area Information Extraction, Entity Linking and (Linked) Open Data, which got merged. On the other hand, at this edition the conference implemented a considerable reduction on the number of area chairs, moving from 30 area chairs in 2016, with two or three area chairs per area, to 16 area chairs in 2017, with one or two area chairs per area, on the basis of the expected number of submissions. On a retrospect, the upper bound of two area chairs per area proved to be a reasonable one, given that the most populated area (Language Resources) received 13 submissions.

In Table 1 we show an overview of all of the thematic areas at CLiC-it 2017, along with the number of accepted papers and their proportion to the whole.

The most successful thematic area, in terms of number of accepted papers, has been Language Resources. Several novel datasets were presented, involving also languages other than Italian, such as English, German and Latin. Two other successful areas are Pragmatics and Creativity, which has covered work on social media, gender analysis, hate speech, irony detection and modality, and the area of Linguistic Issues, which has covered work on lexical semantics, idioms, phrase structure and syntactic selection. Among the remaining areas, very interesting work on deep learning for natural language processing has been presented in Machine Learning.

The Web and Social Media area was characterised mostly by works on affective computing, reporting on specific phenomena, resource creation, and predictive modelling, such as predictions about the Sanremo music festival competition, or controver-

1 Accademia University Press: `http://www.aaccademia.it/scheda-libro?aaref=1186`;
CEUR Workshop Proceedings, AI*IA Series: `http://ceur-ws.org/Vol-2006/`;
OpenEdition Books: `https://books.openedition.org/aaccademia/2314?lang=it`.

sies on Facebook. Finally several works in Digital Humanities focused on a wide variety of data, including a stylometric analysis of the Talmud, and the diachronic distribution of certain noun classes in Latin.

Overall, the conference provided an intellectually stimulating environment for the exchange of ideas, with a broad range of subjects being investigated, as well as a very lively picture of the Italian community working in the field, with nearly half of the participants being PhD students or else junior researchers. The conference also managed to attract a few papers from private companies, including large industrial groups, attesting growing interest for the field also outside of the Italian academia.

3. Worldwide Computational Linguistics at CLiC-it 2017

While CLiC-it is the conference of the Italian Association for Computational Linguistics, it also aims at achieving an international stand. This year edition of the conference has received considerable attention from the international community, with 21 (29%) submissions showing at least one author affiliated to a foreign institution. This amounts to a total of 40 authors over 186 (21%) affiliated to 11 foreign countries: Croatia, Czech Republic, France, Germany, Netherlands, Romania, Spain, Sweden, Switzerland, United Kingdom, and United States.

Conference keynotes also offer a view over topics the international community is currently working on, as seen by renowned scientists world-wide. At the 2017 edition we were lucky to have three excellent invited speakers, namely Marco Baroni (Facebook Artificial Intelligence Research, France), Yoav Goldberg (Bar-Ilan University, Israel), and Rada Mihalcea (University of Michigan, USA). Section 3.1 offers a brief summary of their contributions.

As an additional insight over state-of-the-art international research, and under the suggestion of AILC, CLiC-it 2017 newly introduced a call for Research Communications: authors of articles published in 2017 at outstanding international venues in computational linguistics were encouraged to submit short abstracts of their work to be presented orally at the conference. In Section 3.2 we further elaborate on this initiative and on its outcome at the 2017 edition.

3.1 Keynotes

Marco Baroni's presentation, titled "Spectacular successes and failures of recurrent neural networks applied to language", touched upon deep learning methods and more specifically on the power and limitations of Recurrent Neural Networks (RNNs) when representing linguistic knowledge. Specifically, Marco showed us how he and his colleagues tried to probe the syntactic abilities of RNNs even in absence of meaningful lexical information. In other words: is an RNN getting some grammatical judgments correctly because it relies on lexical information which is naturally intrinsic in a given sentence, or is it really detecting grammaticality per se? Marco illustrated experiments where specific constructions in four different languages (Italian, English, Hebrew, Russian) were tested for the prediction of long-distance number agreement. Nonce examples of such constructions were created so as to deprive the RNN of natural lexical information to see if it could rely on grammatical clues only.

Results above strong baselines in all tested languages indicate that it is indeed likely that *the RNN is learning abstract grammatical representations from linguistic input* (Gulordava et al. 2018). Surprisingly, though, the RNN doesn't seem as skilled when probed on apparently simpler tasks to do with *systematic compositionality*. For example, in one

experiment we saw an RNN trained on phrases containing various commands featuring expressions like "run", "walk", "turn left", "turn right", "run twice", "turn left and run opposite thrice", "walk after run", but including only a small set composed of "jump" commands ("jump", "jump left", "run and jump", "jump around twice"). The system was tested on all the remaining "jump" commands (jump twice, jump left and run opposite thrice, walk after jump), and results indicate that it wasn't able to generalise, having failed to learn compositionality aspects.

Yoav Goldberg's presentation, titled "Doing Stuff with Long Short Term Memory Networks", also focused on deep learning methods for sequence processing, again considering RNNs and specialized versions of these models such as long short term memory (LSTM) networks, which use gating mechanisms. A broad range of tasks in natural language processing have been discussed on which Yoav and his collaborators have been able to achieve state of the art results. More specifically, Yoav has discussed a special attention mechanism used for bidirectional LSTM, that has achieved outstanding results on dependency parsing. Other tasks that have been discussed are coordination boundary prediction, morphological reinflection, preposition sense disambiguation, text generation, and machine translation. Using Yoav's own words, LSTM are very capable learners achieving strong results, making reviewers happy, and resulting in the publication of many papers.

Yoav's presentation also touched upon more theoretical issues, relating RNN models to both linguistic representation and formal language theory. Viewing RNNs as trainable functions from vectors to a single vector, one idea is to inspect what information is encoded in the produced (continuous) vectors. Going through several experiments, Yoav has shown that LSTM are capable of encoding word order information from sentences, as well as sentence lengths. Yoav has also asked the question of what kind of syntactic patterns can be represented by means of RNNs. Partially related to work presented by Marco Baroni, as discussed above, in (Linzen, Dupoux, and Goldberg 2016) Yoav has shown that agreement can be learned remarkably well in simple cases, without the need of supervision. However, in the presence of hierarchical syntactic constructions such as those obtained with the use of relative clauses, there is a performance degradation in learning of agreement dependencies, and some sort of supervision is required. Finally, Yoav has asked the general question of what LSTM models are capable of learning, in relation to formal devices such as finite state automata and grammars providing hierarchical representations. The adopted methodology for this investigation involves inspection of vector representations of sentences, as before, as well as mapping of RNN states into discrete states, forming a finite automaton abstraction. Yoav has shown through several experiments that RNN are capable of capturing regular patterns and, up to some extent, also self-embedding patterns typical of hierarchical structures. However, in many cases the representation captured by the RNN is much more complex than the actual concept class being learned.

Rada Mihalcea held a talk titled "*Computational Sociolinguistics - An Emerging Partnership*" focused on the interplay and mutual benefit between computational linguistics and social sciences. Achievements of the former are currently the trigger for several studies on community phenomena as emerging from social networks, such as the analysis of demographic information as well as the recognition of social trends and personal traits. On the social sciences side, specific tasks and problems stimulated attention on new phenomena where the role of linguistic information is crucial, such as demographic text analysis (Garimella, Banea, and Mihalcea 2017), deception recognition

and grounded emotions. In the talk, Rada covered the different topics by surveying latest results on each one.

On the level of demographic text analysis, the talk discussed the recognition of variations of word associations, as a way to characterise communities at the gender or geographical level. Word associations are crucial signals of the mental model behind conceptual connections in the human mind. These are important to characterise the ways humans, since their young age, develop core components of their semantic knowledge. Moreover, *demographic-aware word association models* are a strong basis for demographic-aware NLP: while community specific word similarity or text similarity models are relatively close tasks, future stages of this research may well incorporate demographic-aware labelled associations and keyword extraction, useful for advanced information retrieval, as well as personalised dialogue. The keynote talk presented several results: word associations do vary in fact across user communties, and automatic discovery methods are able to derive the same patterns as those elicited during traditional classroom surveys. Finally, demographic-aware models, based on a skip-gram architecture, are shown to outperfom user agnostic models (Garimella, Banea, and Mihalcea 2017). A second task discussed in the talk was *deception recognition*. The discussion focused on the role of machine learning and on the impact of a variety of features for the modeling of the deception in open sources as well as in specific settings, such as the multimodal deception detection in real-life situations. The talk suggested that detection can be triggered also against short texts according to simple linguistic features, such as bag of words or bigrams. Although verbal information provides evidence on which the agreement among humans is highest, multimodality is highly beneficial (improvements over 10-15% increase in accuracy). Linguistic information in fact fruitfully combines with non verbal features, such as facial displays (e.g. eyebrow or lips) or hand gestures (e.g head movements and trajectories). Needless to stress that gender and age prediction in deceptive texts is still a challenging task.

3.2 Research Communications

As already mentioned, articles published in year 2017 at major CL/NLP conferences and journals could be orally presented within a dedicated session at the conference, called Research Communications, in order to enforce dissemination of excellence in research. This was mainly thought bearing in mind bachelor and master students who do not often get the opportunity to travel to major events until later in their career, and could therefore get a glimpse of international research carried out at Italian institutions. Furthermore, because CLiC-it requires the submitted papers to be yet unpublished, such research would not have made its way to the conference towards the standard submissions channels, but we still deemed it important that it'd be presented and discussed.[2]

Out of 7 submissions for the Research Communications special track, 5 excellent works were selected and orally presented. These had previously appeared at major 2017 conferences, such as ACL (Croce et al. 2017), EMNLP (Basile and Tamburini 2017), and EACL (Karoui et al. 2017), or had just been published in relevant journals, namely Computational Linguistics (Tripodi and Pelillo 2017) and the ACM Transactions on Interactive Intelligent Systems (Zanzotto and Ferrone 2017).

2 Research communications are not published in the conference proceedings.

4. Students and Junior Researchers

AILC's and CLiC-it's attention to students did not end with having a rich, international, technical programme that young researchers can benefit from. Other initiatives have been put in place with students in mind. Three of them were new at the 2017 edition, while one is by now an established tradition at CLiC-it conferences. The three innovations are the following. First, we organised two tutorials aimed at covering both linguistics and as well more computational aspects, given by international experts in the field (Section 4.1). Second, a panel completely dedicated to the discussion of teaching computational linguistics subjects and programmes in Italy, both at the bachelor and master level, with an eye to Europe (Section 4.2). Third, we introduced a prize for the best thesis in computational linguistics defended in the previous year at any Italian university (Section 4.3). While the panel definitely set and opened a discussion that we believe it will be ongoing in the community but not necessarily represented at future CLiC-it conferences in such a form, we do hope that the tutorials and the prize will become a core part of the annual AILC conferences from this edition onwards.

Finally, as in previous editions of the conference, papers featuring a young researcher as first author could be nominated for the *Young Researcher Best Paper Award* (Section 4.4).

4.1 Tutorials

For the first time in the history of the event, CLiC-it 2017 featured two tutorials, one at the beginning and one at the end of the conference.

The first tutorial, titled "Stretching the Meaning of Words: Inputs for Lexical Resources and Lexical Semantic Models", was provided by Elisabetta Ježek, University of Pavia.[3] This tutorial targeted those researchers in CL/NLP who might be less accustomed to lexical theories. It provided an overview of the main properties of words and a description of the structure of the lexicon in terms of word types, word classes, and word relations. The tutorial also introduced the categories that are needed to classify the types of meaning variation that words display in composition, and examined the interconnection between these variations with syntax, cognition and pragmatics.

The second tutorial, titled "Implementing dynamic neural networks for language with DyNet", was provided by Yoav Goldberg, Bar Ilan University.[4] The tutorial targeted those researchers who want to catch up with state-of-the-art neural approaches, with an applied flavour. Several software libraries are available for programming neural network models, such as Theano, TensorFlow and Keras, which assume a fixed (static) graph structure. This tutorial introduced a radically different approach, the DyNet neural networks package, in which the graphs are dynamic and constructed from scratch for every training example. This makes it very easy to program complex networks with structure that depends on the input. In contrast to existing software, which is tailored for the GPU, the DyNet package also works very well on the CPU.

To particularly highlight the importance that such opportunities have for young researchers, all registered students were allowed to freely attend the tutorials. This was

3 http://sag.art.uniroma2.it/clic2017/Jezek2017Clic-itTutorialRome.pdf
4 http://sag.art.uniroma2.it/clic2017/it/2018/01/04/yoav-golbergs-tutorial-materials/

also made possible thanks to the availability of Elisabetta and Yoav, to whom the CLiC-it 2017 chairs as well as the AILC steering committee are particularly grateful.

4.2 Teaching Computational Linguistics and Natural Language Processing in Italy

Computational Linguistics and Natural Language Processing in Italy are usually not taught as dedicated programmes, with the notable exception of the programme in Digital Humanities (Informatica Umanistica) at the University of Pisa. However, as a community, we do believe this should change in the future as the field deserves a proper, clear position in the Italian higher education sphere.

To make things better we need to first know where we stand. Thus, to glean a picture of the current state of things at Italian Universities, we devised two joint initiatives, as the basis for future developments. First, in Spring 2017 we launched a questionnaire aimed at collecting information over all courses taught on Computational Linguistics and Natural Language Processing both at the bachelor and master levels in Italy[5]. A snapshot of the questions asked in the survey (in Italian) is shown in Figure 1. Respondents were asked to fill one questionnaire per course taught.

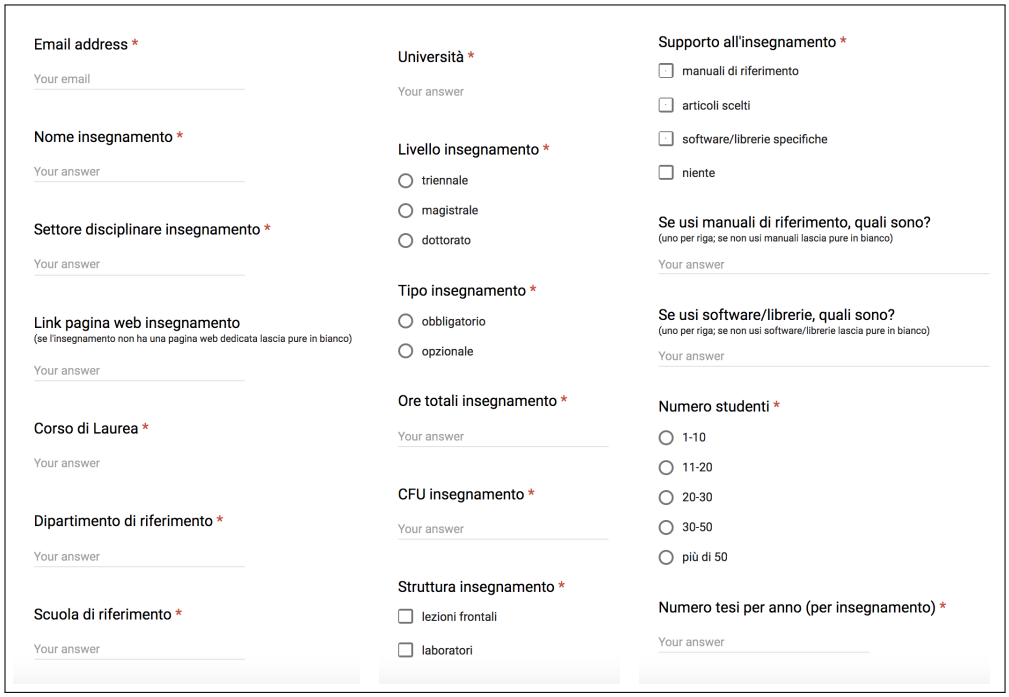

Figure 1
Snapshot of the questionnaire launched in spring 2017 to survey the status of computational linguistics and natural language processing in Italy modules at Italian Universities (`https://goo.gl/9cLzTR`).

Second, we organised a panel at CLiC-it 2017 completely dedicated to teaching. The aim of the panel was to initiate a reflection on the state of things, also discussing the results of the survey. The panel's composition was conceived so as to reflect the

5 `https://goo.gl/9cLzTR`

intrinsic interdisciplinary character of our field, and the fact that courses in computational linguistics are taught within very different programmes. Indeed, we invited a representative of teaching CL/NLP within a humanities programme, a representative of teaching CL/NLP within a science/engineering programme, a representative of teaching CL/NLP within the specific, hybrid Digital Humanities programme in Pisa, and a representative of a gateway to Europe, in the form of the LCT Erasmus Mundus Programme in Language and Communication Technologies. Panelists were therefore as follows:

- Raffaella Bernardi, Università di Trento, Coordinator of the Language and Multimodal Interaction track and local contact of the Erasmus Mundus European Programme in Language and Communication Technologies;
- Alessandro Lenci, Università di Pisa, President of the Degree Programme in Digital Humanities;
- Giovanni Semeraro, Università di Bari "Aldo Moro", Department of Computer Science;
- Fabio Tamburini, Università di Bologna, Department of Classics and Italian Studies (FICLIT).

Each panelist provided a brief overview of their teaching situation and experience, and all together discussed the results of the survey.

At the time of CLiC-it 2017 (December 2017) we had 16 respondents for a total of 26 different courses. The picture that emerges geographically does not accurately reflect the actual situation in Italy, as not everyone responded. Overall, though, the information that we gathered regarding the specifics of the courses is likely to generalise well even to the courses for which no information was provided.

The official areas the courses are taught in are *Linguistics* (L-LIN/01), *Computer engineering* (ING-INF/05), *Computer science* (INF/01), and *Teaching Modern Languages* (L-LIN/02), plus a cross-sector module for PhD students at the University of Trento. The large majority of courses are taught at the master level (77%), and most of the courses overall are optional (61%).

Regarding materials used, there is a wide variety in both fields (humanities and science), with very little overlap. The most used textbook in humanities classes is *Testo e Computer*, by Lenci, Montemagni, Pirrelli (Lenci, Montemagni, and Pirrelli 2005), while the most used in science modules is *Speech and Language Processing*, by Jurafsky and Martin (Jurafsky and Martin 2009). The latter is also one of the only two volumes used in both humanities and science, the other one being *Natural Language Understanding*, by James Allen (Allen 1995). As for tools, those that appear used in both humanities and science modules are OpeNLP, the Stanford Tools, and NLTK, but there is a large number of tools that are only used in computer science modules, and a substantial number of tools used in humanities only. Examples of the former are deep learning modelling frameworks such as Keras, Theano, and Tensorflow, while examples of the latter are more front-end analysis- and annotation-related tools such as the Sketch Engine, the Mate Tools, and Antconc, though we also see tools for morphosyntactic processing, such as TreeTagger and the Malt Parser, in addition to the previously mentioned ones.

The survey and the panel are only just the beginning of an investigation into the Italian teaching context, and while they did open up a reflection and a discussion on the situation of CL/NLP teaching in Italy, the topic will need further and continuous attention. Overall, we believe there is a general consensus in our community over the

need to make our field more independent and more widely recognised as a fully-fledged discipline in the Italian higher education system. Working as a community towards making teaching more homogeneous and more systematically organised is a first step in this direction. The survey, the results and some additional materials are accessible and regularly updated at `https://goo.gl/NZ64Xn`.

4.3 AILC Master Thesis Prize

One more activity intended to recognise excellence in student research was the newly introduced prize for the best Master Thesis (Laurea Magistrale) in Computational Linguistics. This special prize is endorsed by AILC. For this first edition, the committee was composed by a member of the AILC board (Felice Dell'Orletta), a chair of CLiC-it 2016 (Anna Corazza), and a chair of CLiC-it 2017 (Malvina Nissim). Theses defended between January 1st 2016 and July 31st 2017 at any Italian University were eligible for the 2017 prize.

Ten theses were submitted, with the following geographical distribution: Pisa (4),Turin (3), Parma (1), Siena (1), Trento (1). Gender was balanced, with five theses written by female students and five by male students. The evaluation was performed by the three committee members individually in a first stage, after having agreed on a set of specific criteria which had to do both with content (including originality and timeliness of the topic), as well as writing (including clarity, style, and the structure of the thesis). At a second stage, the committee jointly discussed each thesis in details during several Skype meetings, and came up with a short list of three theses, which all deserved the prize. The choice of a final winner was not at all easy, and the reason why eventually we selected the one we selected is its being the closest to the core of our discipline. The first AILC prize for the best master thesis in computational linguistics was thus awarded to:

> Alessio Miaschi, Università di Pisa: "Definizione di modelli computazionali per lo studio dell'evoluzione delle abilità di scrittura a partire da un corpus di produzioni scritte di apprendenti della scuola secondaria di primo grado"

This is a work that involves both the development of a working system that models a specific language phenomenon, as well as a thorough linguistic analysis based on the features used and on detailed error analysis. All this on top of an excellent background overview, and a view to concrete, future applications, directly useful to society.

The other two theses which made it to the final selection were the following:

> Chiara Alzetta, Università di Pisa: "Studio linguistico-computazionale per l'analisi dei tipi linguistici. Similarità e differenze nel confronto fra Universal Dependencies Treebanks"
>
> Enrico Mensa, Università di Torino: "Design and implementation of a methodology for the alignment of semantic resources and the automatic population of Conceptual Spaces"

As part of the prize, Alessio Miaschi received a monetary sum from AILC, free membership to the association for one year, and free attendance to CLiC-it 2017. At the conference the whole community got the chance to listen to Alessio's presentation of his thesis, right at the end of the panel specifically dedicated to the teaching of computational linguistics and Natural Language Processing in Italy. This was a nice

fit, since the high quality of the submitted works really goes to show how much talent, both among students and among teachers, there is at Italian institutions in the field of computational linguistics.

4.4 Young researcher best paper award

In line with previous editions of the conference, CLiC-it 2017 also featured a best paper award, specially directed to PhD students and young researchers. As a short list, the following papers were initially selected by the Program Committee co-chairs on the basis of the review scores and of the above requirement on young authors: "AHyDA: Automatic Hypernym Detection with Feature Augmentation", by Ludovica Pannitto, Lavinia Salicchi and Alessandro Lenci, University of Pisa; "Deep Learning for Automatic Image Captioning in Poor Training Conditions", by Caterina Masotti, Danilo Croce and Roberto Basili, University of Rome Tor Vergata; and "Deep-learning the Ropes: Modeling Idiomaticity with Neural Networks", by Yuri Bizzoni at Göteborg University, Marco Senaldi and Alessandro Lenci at University of Pisa.

In a second phase, a dedicated jury of five people scrutinized and compared the above papers, and took a final decision to assign the award to the paper

> "AHyDA: Automatic Hypernym Detection with Feature Augmentation", by Ludovica Pannitto, Lavinia Salicchi and Alessandro Lenci, University of Pisa

The awarded paper was presented in the area Semantics and Knowledge Acquisition, and reports experiments on a new method of hypernym detection based on a smoothed version of the distributional inclusion hypothesis (Pannitto, Salicchi, and Lenci 2017).

5. Outreach

At its fourth edition, with the fifth one in preparation, CLiC-it has been the prime forum for researchers in Computational Linguistics in Italy. However, the conference also strives to be a platform for discussion on CL/NLP topics also *outside* the research community. We discuss in this section how CLiC-it 2017 has indeed successfully served as a meeting point for researchers, industry, and the public administration.

As Chairs, we organised one panel which revolved uniquely around the work that was done during 2017 by AGID (Agenzia Italia Digitale) through the creation of a dedicated task force on Artificial Intelligence. Launched by the Italian government, this task force has a specific focus on social challenges, opportunities and perspective regulations of AI. The invited panelists were members of the AI task force, including prof. Giuseppe Attardi, Guido Vetere and Enzo Maria Le Fevre, the person responsible for the task force's communication activities. The panel was moderated by Bernardo Magnini, the president of AILC. The discussion has reaffirmed the important role played by Natural Language technologies in the development of the entire AI field. The strong international grounding of Italian research on CL/NLP confirms the potential of the CLiC-it community as a key actor in the task force's activities.

As already mentioned, the conference has received several submissions from private companies, attesting the growing interest for the field also outside of academia. Beside Facebook and other international organisations such as the European Language Resources Association (ELRA), the support to CLiC-it 2018 mainly came from small and medium Italian enterprises, whose core activity is technological innovation, in particular language and voice technologies. The interesting aspect is that most of them

presented also applied research work on specific and somehow innovative problems. Examples in the technical programme are papers on applications in the area of the Italian Public Administration, the adoption of pragmatics cues to improve dialogue abilities in chatbots, the industrial applications of community question answering methods as well as the automatic evaluation of employee satisfaction through the use of written texts and questionnaires. Along with the obvious application-oriented side effects corresponding to effective methods for original applications, these papers confirm the fruitful cross-fertilisation between industrial topics or challenges and the novel paradigms or techniques emerging from academic research.

Finally, another AILC initiative aimed not only at the promotion and development of tools and resources for Italian NLP, but also at collaboration with industry—both in terms of research as well as in terms of end users—is EVALITA, the Campaign for the Evaluation of NLP and speech tools for Italian (`www.evalita.it`). EVALITA is co-located with CLiC-it, but it's a bi-annual event, and was not scheduled for 2017. Next edition will be co-located with CLiC-it 2018. Nevertheless, several papers presented at CLiC-it 2017 made use of the data produced in the context of the 2016 edition of the campaign (Basile et al. 2017), which goes to show the benefits of creating re-usable resources which become benchmarks for a variety of tasks. We are thus very much looking forward to next EVALITA (`www.evalita.it/2018`) and obviously to next CLiC-it in Turin (`clic2018.di.unito.it/it/home`).

Acknowledgments

Even if CLiC-it is a medium size conference, pulling together the meeting requires major efforts on the part of many people. The Program Committee co-chairs would therefore like to take the opportunity to acknowledge here all of the people that have been involved in the event organization. This conference would not have been possible without the dedication, devotion and hard work of the members of the Local Organising Committee, who volunteered their time and energies to contribute to the success of the event. We are also extremely grateful to the Programme Committee members for producing 207 detailed and insightful reviews, as well as to the Area Chairs who assisted us in many ways. We also want to acknowledge the support from endorsing organisations and institutions and from all of the sponsors, who generously provided funds and services that have been crucial for the realisation of the event. Special thanks are also due to the University of Rome "Tor Vergata" and to the National Research Council of Italy for their support in the organisation of the event and for hosting the conference. Finally, we want to acknowledge the EasyChair infrastructure for the management of the review process and the support in the collection of the camera-ready papers for the composition of the conference proceedings.

References

Allen, James. 1995. *Natural Language Understanding (2Nd Ed.)*. Benjamin-Cummings Publishing Co., Inc., Redwood City, CA, USA.

Basile, Ivano and Fabio Tamburini. 2017. Towards quantum language models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 2017)*, Copenhagen, Denmark, September 7-11.

Basile, Pierpaolo, Malvina Nissim, Viviana Patti, Rachele Sprugnoli, and Francesco Cutugno. 2017. EVALITA Goes Social: Tasks, Data. *Italian Journal of Computational Linguistics*, 3(1):93–127.

Croce, Danilo, Simone Filice, Giuseppe Castellucci, and Roberto Basili. 2017. Deep learning in semantic kernel spaces. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 345–354, Vancouver,

Canada, July 30 - August 4.
Garimella, Aparna, Carmen Banea, and Rada Mihalcea. 2017. Demographic-aware word associations. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, (EMNLP 2017)*, pages 2285–2295, Copenhagen, Denmark, September 9-11.
Gulordava, Kristina, Piotr Bojanowski, Edouard Grave, Tal Linzen, and Marco Baroni. 2018. Colorless green recurrent networks dream hierarchically. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1195–1205, New Orleans, Louisiana (USA), June 1-6. Association for Computational Linguistics.
Jurafsky, Daniel and James H. Martin. 2009. *Speech and Language Processing (2Nd Edition)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
Karoui, Jihen, Farah Benamara, Véronique Moriceau, Viviana Patti, Cristina Bosco, and Nathalie Aussenac-Gilles. 2017. Exploring the impact of pragmatic phenomena on irony detection in tweets: A multilingual corpus study. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*. Association for Computational Linguistics (ACL).
Lenci, Alessandro, Simonetta Montemagni, and Vito Pirrelli. 2005. *Testo e computer: elementi di linguistica computazionale*. Studi superiori. Carocci.
Linzen, Tal, Emmanuel Dupoux, and Yoav Goldberg. 2016. Assessing the ability of LSTMs to learn syntax-sensitive dependencies. *Transactions of the Association for Computational Linguistics*, 4:521–535.
Pannitto, Ludovica, Lavinia Salicchi, and Alessandro Lenci. 2017. Ahyda: Automatic hypernym detection with feature augmentation. In *CLiC-it*, volume 2006 of *CEUR Workshop Proceedings*, Roma, December 11-13. CEUR-WS.org.
Tripodi, Rocco and Marcello Pelillo. 2017. A game-theoretic approach to word sense disambiguation. *Computational Linguistics*, 43(1):31–70.
Zanzotto, Fabio Massimo and Lorenzo Ferrone. 2017. Have you lost the thread? Discovering ongoing conversations in scattered dialog blocks. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(2):9.