

# Abstract Brochure

Abstracts for the LCT Annual Meeting Poster Session

June 24–25, 2021

*This brochure contains abstracts for posters being presented at the LCT annual meeting of 2021, hosted virtually by the University of Groningen. The posters are presented in a gather.town virtual poster room, and the section numbers in this brochure correspond to the respective poster slots in the poster session.*



With the support of the  
Erasmus+ Programme  
of the European Union



## Contents

1	Visually Grounded Follow-up Questions: a Dataset of Spatial Questions which Require Dialogue History	3
2	Backtranslation Feedback Improves User Confidence in MT, Not Quality	4
3	Speaking informatively: investigating rational communication strategies in visual contexts	5
4	Visually Grounded Relation Classification of English Noun-Noun Compounds	6
5	Laughing with BERT: Humour Detection using BERT Transformer	7
6	Quality Check and Expansion of Small Treebanks	9
7	Voynich Manuscript Decipherment Approach	11
8	Clause Final Verb Prediction in Hindi: Evidence for Noisy Channel Model of Communication	12
9	A Semantics-Aware Approach to Automated Claim Verification	13
10	Orthography Standardization in Arabic Dialects	14

# 1 Visually Grounded Follow-up Questions: a Dataset of Spatial Questions which Require Dialogue History

**Author:** Tinai Dong (Dota)

**Presented by:** Tinai Dong (Dota)

## **Abstract**

In this paper, we define and evaluate a methodology for extracting history-dependent spatial questions from visual dialogues. We say that a question is history-dependent if it requires (parts of) their dialogue history to be interpreted. We argue that some kinds of visual questions define a context upon which a follow-up spatial question relies. We call the question that restricts the context: trigger, and we call the spatial question that requires the trigger question to be answered: zoomer. We automatically extract different trigger and zoomer pairs based on the visual property that the questions rely on (e.g. color, number). We manually annotate the automatically extracted trigger and zoomer pairs to verify which zoomers require their trigger. We implement a simple baseline architecture based on a SOTA multimodal encoder. It leaves much room for improvement for answering history-dependent questions.

## 2 Backtranslation Feedback Improves User Confidence in MT, Not Quality

**Authors:** Vilém Zouhar, Michal Novák, Matúš Žilinec, Ondřej Bojar, Mateo Obregón, Robin L. Hill, Frédéric Blain, Marina Fomicheva, Lucia Specia and Lisa Yankovskaya

**Presented by:** Vilém Zouhar

### Abstract

Translating text into a language unknown to the text’s author, dubbed outbound translation, is a modern need for which the user experience has significant room for improvement, beyond the basic machine translation facility. We demonstrate this by showing three ways in which user confidence in the outbound translation, as well as its overall final quality, can be affected: backward translation, quality estimation (with alignment) and source paraphrasing. In this paper, we describe an experiment on outbound translation from English to Czech and Estonian. We examine the effects of each proposed feedback module and further focus on how the quality of machine translation systems influence these findings and the user perception of success. We show that backward translation feedback has a mixed effect on the whole process: it increases user confidence in the produced translation, but not the objective quality.

### **3 Speaking informatively: investigating rational communication strategies in visual contexts**

**Authors:** Kathryn V. G. Fisher, Torsten K. Jachmann, Noortje J. Venhuizen and Matthew W. Crocker

**Presented by:** Kathryn V. G. Fisher

#### **Abstract**

In this study we investigate the role informativity plays in the rational communication strategies of speakers in visually-situated contexts. An informative speaker is a fundamental assumption to the formalisations of the Rational Speech Act (RSA, Frank and Goodman, 2012) but has only been realised when viewing informativity in terms of ambiguity or non-ambiguity. To push the assumption of an informative speaker we seek to determine whether speakers are sensitive to graded differences in the informativity of a pair of possible referring expressions. Using a one-shot referential language game, we will examine production choices in response to a visual scene and test whether speakers conform to the predictions of the simplified speaker (S1) model assumed by RSA. If the empirical results do conform to the predictions of the speaker model then this would indicate that speakers do reason about the informativity of an expression. If the results do not conform to the predictions of the speaker model then this would potentially undermine the view of a Gricean informative speaker.

## 4 Visually Grounded Relation Classification of English Noun-Noun Compounds

**Author:** Inga Lang

**Presented by:** Inga Lang

### **Abstract**

Noun-noun compounds (NNCs), i.e. two or more nouns that are put together to form one lexical unit, are extremely frequent and a highly productive category in English, making the interpretation of such phrases crucial for the advancement of several natural language processing tasks, such as text summarization, machine translation, and information retrieval. When interpreting the relation between constituent words in noun-noun compounds, humans utilize experiential knowledge to combine two concepts into one. However, current approaches to automatic noun-noun compound interpretation still utilize sensorily impoverished systems that do not have access to the experiential knowledge humans utilize for this task. We explore the use of visual information for automatic noun-noun compound interpretation in English through a classification task on a large dataset of noun-noun compounds relations built upon compound taxonomies from linguistic theory. Informed and inspired by psycholinguistic research on compound processing, we take a novel approach by introducing visually grounded representations to the task of classifying noun-noun compounds constituent relations and comparing the effect of this in 'traditional' compositional models versus recent deep learning models.

## 5 Laughing with BERT: Humour Detection using BERT Transformer

**Authors:** Alex Lucassen, Sepideh Mollanorozy and Ioanna Tsani

**Presented by:** Alex Lucassen, Sepideh Mollanorozy and Ioanna Tsani

### Abstract

Humour is an integral part of our everyday life and speech. However, the task of automatically detecting it in written form is a very challenging one in Natural Language Processing. In this paper, we attempted to implement a humour detection algorithm using pretrained BERT Transformer on a dataset consisting of one-liner humorous jokes and non-humorous sentences. After experimenting with the parameters of our BERT Transformer model, the system attained an accuracy score of 96.5%, vastly outperforming our baseline system, a feedforward neural network, which only managed to perform the task with an accuracy of 67.8%. The results are very promising, as well as indicative of the fact that humour detection algorithms can be very effective when implemented with the aid of powerful systems such as BERT.

### References

- [1] Issa Annamoradnejad and Gohar Zoghi. Colbert: Using bert sentence embedding for humor detection. *arXiv preprint arXiv:2004.12765*, 2020.
- [2] Yufeng Diao, Hongfei Lin, Liang Yang, Xiaochao Fan, Di Wu, and Kan Xu. Crga: Homographic pun detection with a contextualized-representation: Gated attention network. *Knowledge-Based Systems*, 195:105056, 2020.
- [3] Vaibhav Mathur, Anshu Mathur, and Sunil Kumar. A comparative study of soft computing paradigms for automatic humour detection

- in tweets. In *2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 975–979. IEEE, 2019.
- [4] Tristan Miller and Mladen Turković. Towards the automatic detection and identification of english puns. *The European Journal of Humour Research*, 4(1):59–75, 2016.
- [5] Rida Miraj and Masaki Aono. Integrating extracted information from bert and multiple embedding methods with the deep neural network for humour detection. *arXiv preprint arXiv:2105.05112*, 2021.
- [6] Reynier Ortega-Bueno, Carlos E Muniz-Cuza, José E Medina Pagola, and Paolo Rosso. Uo upv: Deep linguistic humor detection in spanish social media. In *Proceedings of the Third Workshop on Evaluation of Human Language Technologies for Iberian Languages (IberEval 2018) co-located with 34th Conference of the Spanish Society for Natural Language Processing (SEPLN 2018)*, pages 204–213, 2018.
- [7] Yichao Zhou, Jyun-Yu Jiang, Jieyu Zhao, Kai-Wei Chang, and Wei Wang. ” the boating store had its best sail ever”: Pronunciation-attentive contextualized pun recognition. *arXiv preprint arXiv:2004.14457*, 2020.

## 6 Quality Check and Expansion of Small Treebanks

**Author:** Akshay Aggarwal

**Presented by:** Akshay Aggarwal

### Abstract

The Universal Dependencies (UD) initiative, currently involving 183 treebanks [8], has prompted numerous multilingual studies [9, 10] thanks to the large number of gold resources annotated at the morpho-syntactic level under a shared annotation schema. While such studies provide useful linguistic evidence for use in NLP tools [11], only large gold-standards provide enough examples to allow generalisation of linguistic phenomena. Enlargement of small treebanks, constituting  $\approx 60\%$  of UD treebanks<sup>1</sup>, is a crucial challenge second only to treebank quality check, which guarantees that similar constructions share the same annotation.

The present contribution reports part of a study aimed at tackling the two tasks, i.e. treebank *expansion* and *quality check*, in a single workflow. The study exploits LISCA [12], an algorithm proven effective for both tasks [13, 14], integrated into a novel workflow. We use LISCA to obtain a statistical model containing information about the distribution of linguistic phenomena of languages; then, we use the model to identify incorrect analyses from gold treebanks (quality check) and to automatically add examples of under-represented constructions (expansion).

To validate our methodology, we perform experiments on both medium-sized treebanks and small PUD treebanks [10]. In the latter scenario, we utilise the method proposed by [Chapter 6]thesisuni. The results are used to (i) check the inter and intra-treebank homogeneity of PUDs, and (ii) automatically enlarge small treebanks to create richer and more robust sets of examples of language behaviour.

Our experiments suggest that the LISCA-based method can be extended to the case of small treebanks for both quality check and expansion. We believe that such languages would highly benefit from our

---

<sup>1</sup>Word count < 100,000

approach considering that quality control is crucial when dealing with a small amount of data and automatically enlarging resources is valuable for improving their representativeness.

## References

- [8] Daniel Zeman, Joakim Nivre, et al. Universal Dependencies 2.7, 2020. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.
- [9] William Croft, Dawn Nordquist, Katherine Looney, and Michael Regan. Linguistic Typology meets Universal Dependencies. In *Proceedings of TLT15, USA*. CEUR-WS.org, 2017.
- [10] Daniel Zeman, Jan Hajič, Martin Popel, Martin Potthast, et al. CoNLL 2018 shared task: Multilingual Parsing from Raw Text to Universal Dependencies. In *Proceedings of the CoNLL 2018 Shared Task*, Belgium, 2018. ACL.
- [11] Edoardo Maria Ponti, Helen O’Horan, Yevgeni Berzak, Ivan Vulić, et al. Modeling Language Variation and Universals: A Survey on Typological Linguistics for Natural Language Processing. *Computational Linguistics*, 45(3):559–601, 2019.
- [12] Felice Dell’Orletta, Giulia Venturi, and Simonetta Montemagni. Linguistically-driven Selection of Correct Arcs for Dependency Parsing. *Computación y Sistemas*, 17(2):125–136, 2013.
- [13] Chiara Alzetta, Felice Dell’Orletta, Simonetta Montemagni, and Giulia Venturi. Dangerous Relations in Dependency Treebanks. In *Proceedings of TLT16*, Czech Republic, 2017.
- [14] Chiara Alzetta, Felice Dell’Orletta, Simonetta Montemagni, and Giulia Venturi. Uncovering typological context-sensitive features. In *Proceedings of SYGTYP 2020*. ACL, 2020.

## 7 Voynich Manuscript Decipherment Approach

**Authors:** Rasul Jasir Dent, Aditya Kurniawan, Jelena Sarajlić and Evgeniya Ustinova

**Presented by:** Rasul Jasir Dent, Aditya Kurniawan, Jelena Sarajlić and Evgeniya Ustinova

### Abstract

The Voynich Manuscript is an ancient manuscript carbon dated to the 15th century. It is written in an unknown script, and both the authorship and the content are widely debated. There are multiple theories about the contents of Voynich and whether it is a natural language written in its original form, a natural language that has been encrypted, an artificial language or simply - meaningless scribbles. Many researchers have tried to decipher the script, but none succeeded. Our approach uses a method presented in Luo et al. (2021), which uses phonetic prior to decipher ancient languages by searching for cognates between the known and unknown languages. Our assumptions are that "Voynichese" is a known language, just written in an unknown script, and also that the script is alphabetical. The languages of interest we chose to compare to Voynich are Arabic, Italian, Latin, Korean, Maltese, Spanish, and Portuguese. Before starting (to attempt) to decipher Voynich, we test the approach using language pairs of two known languages – Croatian, Serbian and Russian – with the assumption that the model should be able to recognize that Croatian and Serbian are more similar (i.e., there are more cognates found) to each other than Serbian and Russian. Every language of interest is preprocessed by cleaning out the data, stemming or lemmatizing it, and finally transcribing it to IPA symbols, while the unknown language is left unprocessed.

## 8 Clause Final Verb Prediction in Hindi: Evidence for Noisy Channel Model of Communication

**Authors:** Kartik Sharma, Niyati Bafna and Samar Husain

**Presented by:** Niyati Bafna

### **Abstract**

Verbal prediction has been shown to be critical during online comprehension of Subject-Object-Verb (SOV) languages. In this work we present three computational models to predict clause final verbs in Hindi given its prior arguments. The models differ in their use of prior context during the prediction process – the context is either noisy or noise-free. Model predictions are compared with the sentence completion data obtained from Hindi native speakers. Results show that models that assume noisy context outperform the noise-free model. In particular, a lossy context model that assumes prior context to be affected by predictability and recency captures the distribution of the predicted verb class and error sources best. The success of the predictability-recency lossy context model is consistent with the noisy channel hypothesis for sentence comprehension and supports the idea that the reconstruction of the context during prediction is driven by prior linguistic exposure. These results also shed light on the nature of the noise that affects the reconstruction process. Overall the results pose a challenge to the adaptability hypothesis that assumes use of noise-free preverbal context for robust verbal prediction.

## 9 A Semantics-Aware Approach to Automated Claim Verification

**Author:** Blanca Calvo Figueras

**Presented by:** Blanca Calvo Figueras

### **Abstract**

Fake news have become a mainstream topic in the last years due to the fast propagation of misleading information, which has been enhanced by social media. To fight misinformation, researchers have proposed to develop automated solutions. The task of automated claim verification consists in assessing the truthfulness of a claim by finding evidence about its veracity. Datasets with synthetic simple claims have been developed to train models that perform this task. However, naturally-occurring claims and evidences are usually semantically-complex. In this work, we test if the use of explicit semantic structures can help with the task of claim verification. We introduce Semantic Role Labels and Open Information Extraction structures to a BERT model, showing some improvement on the performance of the task. Additionally, we perform explainability tests to assess the linguistic capabilities of our best model.

## 10 Orthography Standardization in Arabic Dialects

**Author:** Christian Cayralat

**Presented by:** Christian Cayralat

### **Abstract**

The Arab world enjoys a wide array of dialects, which are the non-standard varieties of Arabic natively spoken and written on social media across the Arab world. One great obstacle which poses itself in the way of Arabic NLP applications is that Dialectal Arabic lacks a standard orthography system since it is mainly spoken, in addition to being a low-resource language. Because creating more noise-free, manually standardized corpora of Dialectal Arabic does not suffice, since real-world data will most of the time come in a noisy format, this thesis investigates ways to ease the amount of noise in Lebanese Arabic text. It also strives to gain a better understanding about the nature of the noise and its distribution. This is done by leveraging various neural models and by annotating a Lebanese Arabic treebank for spontaneous orthography standardization and morphological segmentation.