**CLASP Annual Report 2020**

Stergios Chatzikyriakidis and Shalom Lappin

February 17, 2021

**1 Introduction**

The covid 19 pandemic has caused most CLASP activities in 2020 to go online. This has prevented us from holding in person seminars, hosting scientific visitors, and conducting conferences on site. However, CLASP staff and students have remained research active, and we have sustained a high level of productive activity. We have appointed a postdoc, and we graduated two additional PhD students, Wafia Adouane and Mehdi Ghanimifard. We organised PaM, an international conference on Probability and Meaning, in October 2020. PaM was sponsored by SIGSEM, and its proceedings have been published in the ACL Anthology for 2020. Our members have published a significant number of papers in high profile international conferences, and they have given invited online talks in a wide range of venues. Our five research groups continue to be active through online reading groups, and joint research activities. We are hoping that 2021 will see a return to some semblance of normality, which will allow us to renew in person meetings.

Shalom Lappin was interviewed on Zoom by the International Panel of the VR, on September 23, 2020, as part of the VR’s mid-term assessment of its international recruitment projects. Stergios Chatzikyriakidis also attended the meeting. Subsequent communication from the VR indicated that the CLASP budget will remain intact. Moreover, on December 10 the VR informed us that they had approved our request to retain the over estimate identified in our original budget for PhD students. This will permit us to appoint a thirteenth PhD student in 2021, and we have now advertised this position.

This is the fifth annual report for the Centre for Linguistic Theory and Studies in Probability (CLASP), covering our activities in 2020. We are grateful to the members of our Scientific Advisory Board, Alex Clark, Katrin Erk, and Joakim Nivre, for invaluable feedback and advice on previous reports. In the current report we cover

(i) staff and recruitment,

(ii) research activities, and our research groups,

(iii) international connections and collaborations,

(iv) scientific activities (workshops, conferences, and visiting

speakers),

(v) PhD program in computational linguistics,

(vi) our alumni

(vii) budget,

(viii) plans for 2021,

(ix) appendix with our publications

(x) appendix with our budget spreadsheet

Additional information on CLASP’s staff, students, recruitment, and regular activities is available on our new external website at https://gu-clasp.github.io/.

**2 CLASP Staff and Recruitment**

In 2020, we hired Julian Grove as a postdoc.

As of December 2020, CLASP was comprised of the following staff:

**Director of CLASP**

Shalom Lappin, Professor

**Associate Director**

Stergios Chatzikyriakidis, Researcher and Associate Director (Docent in Computational Linguistics)

**Administrators**

Iinis Turunen

Susanna Myyry (parental leave – December 2018 – January 2020)

**Researchers**

Robin Cooper, Professor Emeritus, Senior Researcher

Ellen Breitholtz, Researcher

Simon Dobnik, Professor in Computational Linguistics (Docent)

Julian Grove, Postdoctoral Researcher

Christine Howes, Senior Lecturer in Computational Linguistics (Docent)

Staffan Larsson, Professor in Computational Linguistics (Docent)

Sharid Loáiciga Postdocal Researcher (now a Researcher at the University of Potsdam)

Jean-Philippe Bernardy, Researcher (Docent in Computer Science)

Rasmus Blanck, Postdoc (now a Senior Lecturer in logic at FLoV)

Aleksandre Maskharashvili, Postdoctoral Researcher (now a researcher at Ohio State

University)

Asad Sayeed, Senior Lecturer in Computational Linguistics

**Doctoral Students in 2019**

Wafia Adouane (successfully defended her thesis in September 2020)

Adam Ek

Mehdi Ghanimifard (successfully defended his thesis in May 2020)

Kathrein Abu Kwaik

Vladislav Maraev

Bill Noble

Sylvie Saget

Vidya Somashekarappa

Axel Almquist

Tewordros Gebreselassie

Nikolai Ilinykh

**Programmer**

Aram Karimi

**3 Research Activities and Research Groups**

**Dialogue Group:**

Research in the dialogue research group focussed in 2020 on existing projects. These are IncReD and DRiPS (http://www.christinehowes.com/research) which focus on reasoning in dialogue, and GRIPES (https://www.gu.se/en/research/gripes-gothenburg-research-initiative-for-politically-emergent-systems) which looks at the phenomena of dogwhistles, with other members of CLASP and colleagues in the department of Journalism, Media and Communication. Non-project specific research continues in the areas of semantic coordination, dialogue systems, humour, common ground and gaze. We have collaborated extensively both locally and internationally (see publications appendix), and continue to attend internationally acclaimed conferences and workshops to disseminate our work and encourage new collaborations, albeit virtually. In October 2020, we co-hosted CLASP's probability and meaning conference (PaM2020 -- see https://sites.google.com/view/pam2020/home), and published the proceedings via the ACL anthology.

In November 2020 we welcomed Professor Eleni Gregoromichelaki to Linguistics at FLoV, which has already led to a RJ grant application in collaboration with members of the dialogue research group. Other grant applications are in progress to VR and the ERC. The dialogue reading group continues to meet fortnightly via zoom, and break out groups from the core reading group have also led to ongoing research in humour (with a paper currently under revision for Frontiers) and child language acquisition (including a paper at SemDial and a grant proposal under development).

In 2020, several of the invited speakers were specifically relevant to research in the dialogue group, including Shiri Lev-Ari and Elin Macready visiting us in person before the coronavirus lockdowns, and Matthew Purver virtually afterwards. Several of our PhD students also presented their work in progress at the CLASP seminars, including Vlad Maraev, Bill Noble, and Vidya Somashekarappa.

**Machine Learning Group:**

The machine learning group at CLASP took the opportunity to launch many new initiatives this year, some of them reaching across disciplines and departments. The biggest was the start of the Wallenberg-funded GRIPES project on political dogwhistles, a collaboration between linguists, computational linguists, and media studies researchers led by Asad Sayeed.  For GRIPES, this was the year in which we initiated our data collection efforts as well as started the Herculean theoretical task of aligning concepts from opinion research and linguistics.  In addition, the project hired Gregor Rettenegger, a Ph.D. student in media studies with cross-disciplinary expertise in machine learning and computational text analysis, who works closely with Asad.

An additional new collaboration opportunity opened up with the business school. Axel Almquist's work on argument mining has led to opportunities for collaboration with financial reporting researchers into the degree of persuasiveness of corporate earnings reports. In the process, Asad wrote and submitted (in Jan 2021) a joint application, with business school representatives, for infrastructure funds from RJ for network and computing power to process real-time social media and financial reporting data for long-term benefit to multiple existing, and future projects.

Asad continued his work with Yuval Marton and Bloomberg, and expanded that collaboration to include Smaranda Muresan at Columbia University and a team of Columbia University data science students, working on new architectures for representing the frame semantics of verbs.  Asad and Yuval also provided a new preliminary version of the Rollenwechsel semantic roles corpus to Milena Rabovsky at Potsdam University for use in psycholinguistic modeling.

Asad continued to work with researchers at Saarland University and the Max Planck Institute for Informatics on deep learning approaches to visual storytelling, turning a narrative sequence of images into narrative text, producing state of the art results.

In terms of Ph.D. students, Axel started his project in 2020.  His focus is on applied argument mining, and he took on the gargantuan task of assembling and replicating a number of advanced machine learning techniques that have been applied to limited argument mining tasks.  In keeping with his application-orientation, he also developed a powerful, modular workbench for argument mining experimentation.

Vidya succeeded at the Herculean task of designing a video annotation scheme for investigating the social and referential functions of mutual gaze.  With the assistance of a team of master's students, she created a unique resource of annotated videos that can be used in machine learning tasks.  She also developed her skills in computational video processing.

Asad participated in a collaboration led by Bill with Staffan and Raquel Fernández at the University of Amsterdam on analyzing the social network features of language change via NLP techniques over discussion forum corpora.

Finally, the machine learning group opened a new Ph.D. seminar series, "It's evaluation's world, we just live in it", a critical look at the way language science and technology research efforts are evaluated.  This seminar was opened to interested members of the international public, and resulted in on-going discussion across institutions about how natural language processing techniques in particular are evaluated relative to the problems they are intended to solve.

**Probabilistic Natural Language Semantics (PNLS) Group:**

We are currently editing a volume of articles for CSLI Publications, that provide an overview of current developments in probabilistic syntax and semantics. The collection will include contributions from members of CLASP, and a range of other researchers working on probabilistic approaches to natural language. We are contributing a co-authored paper that presents our model of probabilistic semantics. The volume will cover a broad range of methods and models, from small scale Bayesian methods with few parameters, to large-scale deep neural networks with millions of parameters. The intended audience is chiefly graduate students in formal linguistics, NLP, AI, and cognitive science, as well as more advanced researchers seeking to broaden their understanding of work in these areas.

**Situated Language, Perception, and Cognitive Systems Group:**

The situated language and perception group continued its work on (i) computational models and meaning representations of language, action and perception (Dobnik, Ghanimifard, Ilinykh), (ii) semantic models of language and perception (Dobnik, Cooper, Larsson, and Lappin), (iii) co-reference in visual dialogue (Dobnik and Loáiciga), (iv) representation learning in language models (Dobnik, Abu Kwaik, Gebreselassie, Ek, Karimi, Chatzikyriakidis), (v) multi-modal dialogue (Dobnik, Maraev, Noble, Saget, and Somashekarappa). To cover both interests, situated language processing and representation learning, the group has also come to be known as a cognitive systems group. For details of publications and activities, please refer to the appendix entries for individual researchers.

In January 2020 three new doctoral students joined CLASP whose research closely aligns with that of the group: Nikolai Ilinykh (vision and language), Tewodros Gebreselassie (representation learning for an under-resourced language), and Axel Almquist (representation learning and argument mining). Mehdi Ghanimifard successfully defended his thesis on May 27th. He continue his collaboration with the group until January this year, mostly for teaching masters courses. Then he started as a data scientist at Seal Software in Gothenburg, which is a part of DocuSign. Sharid Loáiciga is now a member of David Schlangen's group in Potsdam, but she is also continuing her collaboration with Dobnik on coreference in visual dialogue. Between March and October Maryam Rajestari joined the group as a researcher on the Erasmus traineeship grant working on the gradability of (allegedly) non-gradable adjectives.

Locally, the group, which consists of senior researchers and doctoral students, meets fortnightly as a reading group for which students can also receive course credit. (Two doctoral students from Språkbanken are also regular attendees.) Currently, there are three standing doctoral courses: Language, Action, and Perception (APL), Representations of Meaning (ROM) and Machine Learning Methods for Vision and Language (ML-V&L). Additional reading courses can be arranged to cover more specific interests of students.

Dobnik and Larsson took part in the organisation of the Eighth Swedish Language Technology Conference (SLTC), which took place between 25 and 27 November virtually in Gothenburg. At SLTC, Dobnik and Gebreselassie also co-organised (with Barbara Plank and Lars Borin) the RESOURCEFUL workshop on RESOURCEs and representations For Under-resourced Languages and domains. Several members were involved in the organisation of the Probability and Meaning (PaM) CLASP conference.

Group leader Simon Dobnik has continued his work in the Coordination Group for Artificial Intelligence and Digitisation at University of Gothenburg (SAID), which acts as an advisory body to the deputy vice-chancellor on issues related to AI. Here, the main focus was extending collaboration with Chalmers Technical University and other Swedish programmes related to AI: AI Sweden and AI Competence. From these sources funding was obtained to organise an open course or summer school with external lectures on natural language processing, vision, and robotics (which was postponed due to pandemic), and the language technology module of a MOOC "AI and Applications" that is being designed by Gothenburg University.

The group is working on research grant applications to the VR, to expand its activities.

Two events that were planned for 2020 could not take place due to pandemic, and they have been postponed to 2021 or 2022. These are Dobnik, Kelleher and Ghanimifard’s course accepted for NASSLLI-2020, and a CLASP industry day on the relation between AI and NLP (co-organised by Dobnik and Larsson) which is designed to further industrial connections for CLASP members, and to expand career opportunities for masters and doctoral students.

**Type Theory Group**

The Type Theory group is devoted to the study of Type Theoretical methods for NLP and Formal Semantics. The group is led by Stergios Chatzikyriakidis.

Group members:

* Stergios Chatzikyriakidis
* Jean-Philippe Bernardy
* Julian Grove
* Shalom Lappin
* Robin Cooper
* Staffan Larsson
* Vlad Maraev
* Bill Noble
* Ellen Breitholtz

The group’s activities involve invited talks by prominent researchers in the field, organizing workshops on Type Theory and exploring connections between Type Theory and Probability and/or Machine Learning. In 2020, the group has focused on a number of issues ranging from Natural Language Inference using type theoretic methods, combining type theory with effects (monads, applicative functors, continuations), as well as more formal issues using type theories for Natural Language Semantics and/or Computational Semantics

This year, due to the Corona, we did not get the chance to invite core researchers in Type Theory to visit us, the exception being Lasha Abzianidze’s visit on February. We have however managed to have virtual talks by the following people (see the seminar section for titles for the talks:

- Daisuke Bekki

- Herbert Lange

- Wlodek Zadrozny

- Lasha Abzianidze (two talks)

- Julian Grove

Stergios Chatzikyriakidis was part of the organization of the NALOMA workshop at NASSLLI 2020 (weSSLLI2020,<https://typo.uni-konstanz.de/naloma20/)>. The following workshop is scheduled for ESSLLI 2021:

- Stergios Chatzikyriakidis and Rainer Osswald. Workshop on Computing Semantics with Types, Frames and Related Structures. ESSLLI2021 (moved from last year)

The members of the group produced several papers in major conferences, journal articles, and book chapters. These are listed in the Publications Appendix. Jean-Philippe Bernardy, Stergios Chatzikyriakidis and Julian Grove applied for Joint Japan-Sweden Research Collaboration grant (the other site being Daisuke Bekki’s group in Tokyo).

**4 Scientific Activities (Conferences, Workshops, Seminars, and Speakers)**

The First Conference on Probability and Meaning (PaM2020) PaM’s aim was to bring together researchers interested in computationally relevant probabilistic approaches to natural language semantics. PaM included symbolic, machine learning and experimental approaches to this task, as well as hybrid models. The conference moved into a fully online format. More info can be found here:

<https://sites.google.com/view/pam2020/home>. The conference was sponsored by SIGSEM and its proceedings were published at the ACL anthology:

<https://www.aclweb.org/anthology/volumes/2020.pam-1/>

The AI and Natural Language industry day that was originally scheduled to co-locate with PaM was moved to a later date to be specified soon:

<https://sites.google.com/view/pam2020/industry-day>

**5 Joint Research with External Partners**

Ellen Breitholtz, Christine Howes and Robin Cooper continued to work with Mary Lavelle (City, University of London) on the DRiPS project.

Stergios Chatzikyriakidis has continued his longterm collaboration with Zhaohui Luo from Royal Holloway, University of London. This year their monograph Formal Semantics in Modern Type Theories has been published by ISTE/Wiley (see the publications section for more details)

Christine Howes works with Arash Eshghi (Heriot-Watt, Edinburgh) on feedback in dialogue. This work has led to a paper being accepted for publication in the Journal of Logic, Linguistics and Information (to appear in 2021).

Shalom Lappin has been working with members of the NLP lab in the Cognitive Science Group in the School of Electronic Engineering and Computer Science, at Queen Mary University, and with Jey Han Lau of the Department of Computer Science at the University of Melbourne, on using deep learning models to predict human sentence acceptability judgements in context. The results of this joint research were published in a paper in *Transaction of the Association of Computational Linguistics*, 2020. They have also submitted a research grant application to the Engineering and Physical Sciences Research Council UK to fund extensions of this work to prediction of acceptability judgements in extra linguistic contexts.

Asad Sayeed was involved in several international research collaborations, which are described in the above section of the Machine Learning Research Group.

Simon Dobnik has been working with John D Kelleher from Dublin Technical University on research on computational modelling of spatial language and organisation of a summer school course and with Sharid Loáiciga at Potsdam University on coreference resolution in visual dialogue.

**Our Seminars, Visiting Speakers/Guest Researchers**

Despite the pandemic, CLASP continues to be very active in conducting regular seminars in which members of CLASP, visiting speakers, and guest researchers present work related to CLASP’s scientific interests. The number of physical visits has been of course drastically reduced this year and covers only the period of the first three months of 2020 (Shiri Lev-Ari, Lasha Abzianidze and Elin McCready), but the seminar series has been running virtually with a wide range of guest speakers. The topics covered by the seminars indicate CLASP’s broad range of research activities. A full list of CLASP seminars, with titles, abstracts, and slides, is available at:<https://gu-clasp.github.io/event/?type=seminars&page=1>

- Shiri Lev-Ari “Language evolution and change from a social networks perspective”

- Lasha Abzianidze The Parallel Meaning Bank: a corpus of translations annotated with formal meaning representations

- Lasha Abzianidze “Natural theorem proving for natural language: theory and application”

- Desmond Elliott “Compositional Generalization in Image Captioning”

- Elin McCready “Dogwhistles: Ideology and Trust”

- Asad Sayeed and Yuval Marton “Data requirements for thematic fit modelling”

- Herbert Lange “Learning Domain-Specific Grammars from Example Sentences”

- Matthew Purver "Transferable skills: using and comparing contextual embeddings across languages”

- Gemma Boleda “Distributional Semantics and Linguistic Theory”

- Wlodek Zadrozny “Towards Co-Inductive Models for Natural Language Semantics”

- Milena Rabovsky “Modeling the N400 brain potential as change in a probabilistic representation of meaning”

- Emil Carlsson and Devdatt Dubhashi “Learning numeral systems by interaction”

- Ellie Pavlick “You can lead a horse to water...: Representing vs. Using Features in Neural NLP”

- Eleni Gregoromichelaki “Affordances for syntax and semantics”

- Daisuke Bekki “A hybrid approach toward Natural Language Understanding”

**6 PhD Program in Computational Linguistics**

We currently have 9 PhD students in CLASP. They are

1. Kathrein Abu-Kwaik (started November 2016, first supervisor Stergios Chatzikyriakidis, second supervisor Simon Dobnik)

2. Vladislav Maraev (started September 2017, first supervisor Chris Howes, second supervisor Staffan Larsson, Jonathan Ginzburg, external advisor)

3. Sylvie Saget (started September, 2017, first supervisor (until November 2020) Stergios Chatzikyriakidis, second supervisor Dominique Knutsen (Université de Lille))

4. Bill Noble (started September 2018, first supervisor Staffan Larsson, second supervisor Asad Sayeed)

5. Adam Ek (started March 2019, first supervisor Jean-Philippe Bernardy, second supervisor Stergios Chatzikyriakidis)

6. Vidya Somashekarappa (started January 2019, first supervisor Asad Sayeed, second supervisor Chris Howes)

7. Axel Almquist (started January 2020, first supervisor Asad Sayeed, second supervisor Simon Dobnik)

8. Tewordros Gebreselassie (started January 2020, first supervisor Simon Dobnik, second supervisor, Stergios Chatzikyriakidis)

9. Nikolai Ilinykh (stating February 2020, first supervisor Simon Dobnik, second supervisor, Asad Sayeed)

Our students are making excellent progress, with most of them presenting papers at major international conferences over the past year. Wafia Adouane successfully defended her thesis in September 2020 and Mehdi Ghanimifard in May 2020.

**7 Budget**

After our mid-term interview with the VR international panel, the VR indicated to us that they expect our budget to remain intact over the second half of our grant period. They also kindly permitted us to retain the estimated surplus in our PhD student budget, allowing us to recruit an additional student.

Due to the pandemic, international travel has been greatly reduced during 2020. As a result, we have spent little of our budgets for travel, and for scientific visitors. Our conference budget was also largely conserved during this period, as our events have been conducted on Zoom. Therefore, we have run up a substantial budget surplus over the past year. We fully intend to use these funds for CLASP research activities, as restrictions ease in the future.

The accompanying budget spread sheet, prepared by Pia Gårdmo, FLoV’s senior financial officer, shows us having a small surplus at the end of the tenure of the grant in 2025. The projected expenses include the additional PhD student, and the two postdoctoral researchers that we will be appointing over this year and next.

**8 CLASP Alumni**

**Wafia Adouane**

Since completing her PhD at CLASP in 2020, Wafia has been preparing to apply for postdoctoral research positions, and jobs in industrial research labs.

**Yuri Bizzoni**

Since completing his PhD at CLASP in 2019, Yuri has been a postdoctoral researcher in an SFB computational linguistics project at the University of Saarlandes.

He works on two distinct, but related projects within a larger group focused on studying language use through the lens of information theory. One project focuses on the evolution of scientific style. Using the longest standing scientific journal, the Royal Society's, as a corpus, he and his colleagues observe how disciplines develop through time a specific style optimized to convey information in more efficient ways. In the other project, they attempt to design translation as a form of communication through a noisy channel, trying to quantify how much information gets lost or distorted during the process.

**Rasmus Blanck**

Since completing his postdoctoral research position in CLASP in 2020, Rasmus has been teaching in FLoV as a (temporary) senior lecturer in logic and theoretical philosophy.

**Mehdi Ghanimifard**

Since completing his PhD at CLASP in 2020, Mehdi is a Data Scientist at DocuSign. He is working on natural language understanding for contracts and agreement.

**Sharid Loáiciga**

Since completing her postdoc position at CLASP in 2020, Sharid has been a researcher in NLP in David Schlangen's group at the University of Potsdam. She has pursued her research on event pronouns and coreference with several colleagues from different universities. She has also continued work on the manual annotation of a multimodal corpus with coreference information that she had started at CLASP. This annotation work has now evolved from a small pilot into a big ongoing project. Sharid expects to spend most of her time on this work until the summer of 2021. Together with her Potsdam colleagues, she has started evaluation of big pre-trained language models.

Sharid is teaching 4 hours a week, and she is supervising several master students.

Sharid’s publications for last year:

Patrick Kahardipraja, Olena Vyshnevska, Sharid Loáiciga (2020), “Exploring Span

Representations in Neural Coreference Resolution”, *Proceedings of the CODI Workshop*.

Sharid Loáiciga, Christian Hardmeier and Asad Sayeed (2020), “Exploiting Cross-Lingual

Hints to Discover Event Pronouns”, *Proceedings of the12th Language Resources and*

*Evaluation Conference (LREC20)*.

**Alexander Maskharashvili**

Since completing his postdoctoral position at CLASP in 2020, Alexander Maskharashvili has been a researcher in a computational linguistics project at Ohio State University. Hiswork there concerns natural language generation in conversational systems. He has been studying whether structured input (eg. tree/graph representations of discourse involving rhetorical relations) helps neural network based models to generate coherent texts. His group is also trying to understand how large pretrained models (eg. BERT, BART, GPT-2, and T5) can be used to achieve improved performance in NLG. He has co-authored two conference papers on these topics.

Symon Stevens-Guille, Aleksandre Maskharashvili, Amy Isard, Xintong Li, Michael White

(2020), “Neural NLG for Methodius: From RST Meaning Representations to Texts”,

*Proceedings of the 13th International Conference on Natural Language Generation*

*(INLG2020)*, ACL Anthology

Xintong Li, Aleksandre Maskharashvili, Symon Jory Stevens-Guille and Michael White

(2020), “Leveraging Large Pretrained Models for WebNLG 2020”, *WebNLG+ 2020 : 3rd*

*Workshop on Natural Language Generation from the Semantic Web*.

**9 Plans for 2021**

**New Appointments:**

Our Associate Director, Stergios Chatzikyriakidis, has accepted a professorship at the University of Crete, which he will take up in August 2021. Stergios has made an enormous contribution to CLASP. We are sorry to see him go, while, at the same time, we celebrate his success in moving into a senior academic role, which allows him to return home to Greece. We are recruiting candidates to fill his position, and we hope to appoint the successful candidate by May, 2021. This will give us several weeks of overlap between the Stergios and the incoming Associate Director.

We are also recruiting for an additional postdoctoral researcher, and our thirteenth PhD student. We expect to fill these positions by September 2021.

**Annual Conference and Workshops:**

We are now planning our annual Workshop/Conference for October 2021, and we are hoping that it may be possible to conduct this as an in person event at GU. The topic of the conference is reasoning in natural language, and it will feature two associated workshops. One will be devoted to inference in dialogue, and it will have a shared task. Bill Noble and Adam Ek have taken the lead in organising this event and the shared task. The theme of the second workshop is extra-linguistic grounding in interpretation, and Asad Sayeed will be chairing this meeting. We will be approaching SiGSEM about the possibility of sponsorship, and we hope to publish the proceedings of the conference, with the papers from the workshops, as a SiGSEM event, in the ACL Anthology for 2021.

**Scientific Visitors and Participation in Conferences:**

While covid restrictions persist we will continue to conduct our seminars, and to participate in international conferences, online. We are hoping that the situation will improve to the point that we can begin to resume in person visits and meetings in the second half of the year. Fortunately, we have managed to sustain a high degree of productivity in research and publication in 2020, despite the hardships of the pandemic. We plan to continue our research momentum in the coming year.

**10 Appendix: Publications, Talks, and Other Scientific Activities**

**Axel Almquist**

**PhD Thesis Work:**

During 2020 I built a Framework for Argument Mining. By Combining open-source libraries in Machine Learning/NLP/Visualization I have created an easy-to-use python library, where anyone can train and evaluate models for Argument Mining. The aim is to streamline Argument Mining research, making it more accessible and easie, while providing common ground for evaluation.

**Seminar:**

HotAM: An Argument Mining Framework, CLASP Seminar, February 3, 2020.

**Jean-Philippe Bernardy**

**Journal Articles:**

Jean-Philippe Bernardy and Stergios Chatzikyriakidis (2020), “A Computational Treatment of

Anaphora and its Algorithmic Implementation”, *Journal of Logic, Language and*

*Information*.

Andreas Abel and Jean-Philippe Bernardy (2020), “A Unified View of Modalities in Type

Systems”, *Proceedings of the ACM on Programming Languages 4.ICFP*.

**Papers in Refereed Conference Proceedings:**

Adam Ek and Jean-Philippe Bernardy (2020), “Composing Byte-Pair Encodings for

Morphological Sequence Classification”, *Proceedings of the Fourth Workshop on*

*Universal Dependencies*.

Yves Parès, Jean-Philippe Bernardy, and Richard A. Eisenberg (2020), “Algebraic Effects

with Tasks Build Composable Data Workflows”, *Proceedings of the Haskell Symposium*.

Adam Ek and Jean-Philippe Bernardy (2020), “How does Punctuation Affect Neural Models

in Natural Language Inference” *Proceedings of the conference on Probability and*

*Meaning*.

Adam Ek and Jean-Philippe Bernardy (2020), “How much of enhanced UD is contained in

UD?”, *Proceedings of the 16th International Conference on Parsing Technologies and the*

*IWPT 2020*, Shared Task on Parsing into Enhanced Universal Dependencies.

Wafia Adouane and Jean-Philippe Bernardy (2020), “When is Multi-task Learning Beneficial

for Low-Resource Noisy Code-Switched User-generated Algerian Texts?”, *Proceedings of*

*the Fourth Workshop on Computational Approaches to Linguistic Code-Switching*.

Wafia Adouane, Samia Touileb, and Jean-Philippe Bernardy (2020), “Identifying Sentiments

in Algerian Code-switched User-generated Comments”, *Proceedings of the Language*

*Resources and Evaluation Conference*.

Jean-Philippe Bernardy and Stergios Chatzikyriakidis (2020), “Improving the Precision of

Natural Textual Entailment Problem datasets”, *Proceedings of the Language Resources*

*and Evaluation Conference*.

**Non-Peer Reviewed Publication:**

Jean-Philippe Bernardy, Rasmus Blanck, and Aleksandre Maskharashvili (2020), “A Logic

with Measurable Spaces for Natural Language Semantics”, *Applied Mathematics,*

*Informatics and Mechanics*, Special issue dedicated to Shalva Pkhakadze.

**Ellen Breitholtz**

**Monograph:**

Breitholtz, Ellen (2020), *Enthymemes and Topoi in Dialogue: The Use of Common*

*Sense Reasoning in Conversation*, volume 41 of Current Research in the

Semantics/Pragmatics, Brill.

**Book Chapters:**

Breitholtz, Ellen; Cooper, Robin; Howes, Christine; and Lavelle, Mary (2020), “Multiparty

Reasoning Involving Patients with Schizophrenia, in (In)Coherence of Discourse”,

Language Cognition and Mind. Springer.

**Papers in Refereed Conference Proceedings:**

Breitholtz, E. & Howes, C (2020), “Communicable Reasons: How Children Learn Topoi

through Dialogue, *Proceedings of the 24th Workshop on the Semantics and Pragmatics of*

*Dialogue*, Waltham, MA : SEMDIAL.

Breitholtz, E. & Howes, C (2020), “Theirs Not to Reason Why: Dialogical Reasoning for

Conversational Artificial Agents*, Do Robots Talk? Philosophical Implications of*

*Describing Human-Machine Communication (DoRoTa)*.

Maraev, V., Breitholtz, E., and Howes, C. (2020a), "How Do You Make an AI Get the Joke?

Here’s What I Found on the Web", *First AISB Symposium on Conversational AI (SoCAI)*.

Noble, Bill, Ellen Breitholtz and Robin Cooper (2020), Personae under Uncertainty: The case

of Topoi, in Stergios Chatzikyriakidis, Christine Howes (Editors), *Proceedings of the*

*Probability and Meaning Conference (PaM 2020)*, SiGSEM, ACL Anthology.

**Edited Volumes:**

Howes, C., Dobnik, S. & Breitholtz, E., editors (2020), *Dialogue and Perception - Extended*

*Papers from DaP2018*, CLASP papers in computational linguistics, Gothenburg, GUPEA.

**Invited Talk:**

Ellen Breitholtz, A Contemporary Take on Aristotelian Enthymemes and Topoi in Dialogue

at Pragmatic Argumentation in the History of Philosophy, The Swedish Institute in Rome, February 6-7 2020.

**Stergios Chatzikyriakidis**

**Monograph:**

Chatzikyriakidis, S. and Luo, Z. (2020), *Formal Semantics in Modern Type Theories*,

ISTE/Wiley Science.

**Journal Articles:**

Bernardy, J.P and Chatzikyriakidis S. and Maskharashvili A. (2020), “A Computational

Treatment of Anaphora and its Algorithmic Implementation”, *Journal of Language, Logic*

*and Information.*

Chatzikyriakidis, S. (2020), “Underspecification, Parsing Mismatches and Routinization: The

Historical Development of the Clitic Systems of Greek Dialects”, *Journal of Language,*

*Logic and Information*, special issue on Dynamic Syntax, edited by Gregoromichelaki, E.,

Mills, G. Howes.

Gregoromichelaki, E., Mills, G. J., Howes, C., Eshghi, A., Chatzikyriakidis, S., Purver, M., Kempson R., Cann, R. and Healey, P. (2020), “Completability vs (In)completeness”, *Acta Linguistica*

*Hafniensia.*

**Papers in Refereed Conference Proceedings:**

Chatzikyriakidis S. (2020). “Dialogue Informing Syntax/Semantics: The Case of

Afterthoughts”, P*roceedings of SemDial* (WatchDial2020).

Gregoromichelaki, E., Chatzikyriakidis, S., Eshghi, A., Hough, J., Howes, C., Kempson, R.,

Kiaer, J., Purver, M., Sadrzadeh, M. and White, G. (2020), “Affordance Competition in

Dialogue: The Case of Syntactic Unversals”, P*roceedings of SemDial* (WatchDial2020).

​Ek, E., Bernardy, JP and Chatzikyriakidis S. (2020), “How does Punctuation Affect Neural

Models in Natural Language Inference”, *Proceedings of the Conference on Probability and*

*Meaning*, ACL anthology.

Qwaider, C., Saad, M., Chatzikyriakidis, S. and Dobnik S. and Johansson. R. (2020), “An

Arabic Tweets Sentiment Analysis Dataset (ATSAD) Using Distant Supervision and Self

Training”, *Proceedings of the 4th Workshop on Open-Source Arabic Corpora and*

*Processing Tools*, LREC2020.

Bernardy, JP. and Chatzikyriakidis S. (2020), “Improving the Precision of Natural Textual

Entailment Problems”, *Proceedings of LREC2020*, Marseille, France.

**ArXiv paper:**

Bernardy, JP. and Chatzikyriakidis S., 2020. *FraCaS: Temporal Analysis*. ArXiv: 2021.10668.

**Edited Volumes:**

Howes, C., Chatzikyriakidis, S., Ek, A. and Somashekarappa, V. (2020), *Proceedings of the 1st Conference on Probability and Meaning*, SiGSEM, ACL anthology.

**Conference Organising:**

Conference on Probability and Meaning (PaM2020). Co-chair with Christine Howes

Natural Logic meets Machine Learning (weSSLLI2020). Organizing committee.

**Other Scientific Activity:**

CLASP seminar organizer and coordinator

**Robin Cooper**

**Chapters in Volumes:**

Jonathan Ginzburg, Robin Cooper, Julian Hough and David Schlangen (2020),

“Incrementality and HPSG: Why not?” in Anne Abeillé and Olivier Bonami (eds.),

*Constraint-Based Syntax and Semantics: Papers in Honor of Danièle Godard*, CSLI

Publications, Stanford CA.

**Papers in Refereed Conference Proceedings:**

Maryam Rajestari, Simon Dobnik, Robin Cooper and Aram Karimi (2020), "Very

Necessary: the Meaning of Non-Gradable Modal Adjectives in Discourse Contexts",

*Proceedings of the Eighth Swedish Language Technology Conference (SLTC)*, 25–27

November, Gothenburg.

Noble, Bill, Ellen Breitholtz and Robin Cooper (2020), Personae under Uncertainty: The case

of Topoi, in Stergios Chatzikyriakidis, Christine Howes (Editors), *Proceedings of the*

*Probability and Meaning Conference (PaM 2020)*, SiGSEM, ACL Anthology.

**Invited Summer Institute Course:**

(with Staffan Larsson), Modelling Linguistic Communication Using Types, WeSSLLI, Brandeis, 11th-17th July, 2020.

**Simon Dobnik**

**Papers in Refereed Conference Proceedings:**

S. Dobnik and M. Ghanimifard (2020), “Spatial Descriptions on a Functional-Geometric

Spectrum: The Location of Objects”, J. Šķilters, N. S. Newcombe, and D. Uttal, editors,

*Spatial Cognition XII: Proceedings of the 12th International Conference, Spatial Cognition*

*2020*, Riga, Latvia, pages 219–234, Cham, Switzerland, Springer International

Publishing.

J. M. Cano Santín, S. Dobnik, and M. Ghanimifard (2020), “Fast Visual Grounding in

Interaction: Bringing Few-Shot Learning with Neural Networks to an Interactive Robot”,

*Proceedings of Conference on Probability and Meaning (PaM-2020)*, Gothenburg,

Sweden, pages 1–9, Gothenburg, Sweden, October 14–16 2020, Association for

Computational Linguistics (ACL), Special Interest Group on Computational Semantics

(SiGSEM).

J. D. Kelleher and S. Dobnik (2020), “Referring to the Recently Seen: Reference and

Perceptual Memory in Situated Dialogue”, C. Howes, S. Dobnik, and E. Breitholtz, editors,

*CLASP Papers in Computational Linguistics: Dialogue and Perception – Extended papers*

*from DaP-2018 Gothenburg,* volume 2, pages 41–50, Gothenburg, Sweden, February 2020.

University of Gothenburg, CLASP, Centre for Language and Studies in Probability and

GUPEA.

M. Rajestari, S. Dobnik, R. Cooper, and A. Karimi (2020), “Very Necessary: the Meaning of

Non-Gradable Modal Adjectives in Discourse Contexts”, *Proceedings of the Eighth*

*Swedish Language Technology Conference (SLTC)*, pages 1–5, Gothenburg, Sweden, 25–

27 November 2020.

Y. Bizzoni and S. Dobnik (2020), “Sky + Fire = Sunset: Exploring Parallels between Visually

Grounded Metaphors and Image Classifiers”, B. B. Klebanov, E. Shutova, P. Lichtenstein,

S. Muresan, C. Wee, A. Feldman, and D. Ghosh, editors, *Proceedings of the Second*

*Workshop on Figurative Language Processing (FLP)*, pages 126–135, July 2020,

Association for Computational Linguistics.

N. Ilinykh and S. Dobnik (2020), “When an Image Tells a Story: The Role of Visual and

Semantic Information for Generating Paragraph Descriptions”, B. Davis, Y. G. and John

Kelleher, and Y. Sripada, editors, *Proceedings of the 13th International Conference on*

*Natural Language Generation*, pages 338–348, Dublin, Ireland, Dec. 2020, Association for

Computational Linguistics.

K. Abu Kwaik, S. Chatzikyriakidis, S. Dobnik, M. Saad, and R. Johansson (2020), “An

Arabic Tweets Sentiment Analysis Dataset (ATSAD) Using Distant Supervision and Self

Training”, H. Al-Khalifa, W. Magdy, K. Darwish, T. Elsayed, and H. Mubarak, editors,

*Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools*

*with a Shared Task on Offensive Language Detection (OSACT4- 2020)*, Language

Resources and Evaluation Conference (LREC 2020), pages 1–8, Marseille, France, May

11–16 2020. European Language Resources Association (ELRA).

L. Strand, R. Rhys Thomas, S. Dobnik, and M. Ghanimifard (2020), “Topic Modelling for a

Virtual Librarian Assistant Tool”, *Proceedings of WatchDial - Semdial 2020: The 24th*

*Workshop on the Semantics and Pragmatics of Dialogue*, pages 1–4, Waltham, MA, 18–19

July 2020. Brandeis University.

M. Sandahl, S. Dobnik, and M. Ghanimifard (2020), “A Corpus of Swedish Conversations

with a Librarian”, *Proceedings of WatchDial - Semdial 2020: The 24the Workshop on the*

*Semantics and Pragmatics of Dialogue*, pages 1–3, Waltham, MA, 18–19 July 2020.

Brandeis University.

S. Dobnik, J. D. Kelleher, and C. Howes (2020), “Local Alignment of Frame of Reference

Assignment in English and Swedish dialogue”, J. Šķilters, N. S. Newcombe, and D. Uttal,

editors, *Spatial Cognition XII: Proceedings of the 12th International Conference, Spatial*

*Cognition 2020,* Riga, Latvia, pages 251–267, Cham, Switzerland, 2020. Springer

International Publishing.

**Edited Volumes:**

C. Howes, S. Dobnik, and E. Breitholtz (2020), editors, *CLASP Papers in Computational*

*Linguistics: Dialogue and Perception – Extended papers from DaP-2018 Gothenburg*,

volume 2, Gothenburg, Sweden, February 2020. University of Gothenburg, CLASP, Centre

for Language and Studies in Probability and GUPEA.

**Research Grants:**

Faculty funding for seminars of the Centre for Language Technology (CLT)

AI Competence for Sweden, https://ai-competence.se/en/: open masters/PhD course with external lecturers on natural language processing, vision and robotics

Gothenburg University: to co-develop a language technology module of a MOOC "AI and Applications"

**Workshop Organising:**

Local co-organiser and programme co-chair of the Eighth Swedish Language Technology Conference (SLTC), https://spraakbanken.gu.se/en/sltc2020

Co-organiser of the RESOURCEFUL-2020 (RESOURCEs and representations For Under-resourced Languages and domains) workshop at SLTC, https://gu-clasp.github.io/resourceful-2020

**Other Scientific Activity:**

Co-organiser of seminars of the Centre for Language Technology (CLT) at Gothenburg University

Member of the steering committee of the Centre for Digital Humanities/Centrum för Digital Humanioras

Member of the SAID group (Samordningsgrupp för artificial intelligens och digitalisering, Coordination group for artificial intelligence and digitalisation) of University of Gothenburg, https://www.gu.se/ai-digitalisering

Appointed as a Professor of Computational Linguistics

**Adam Ek**

**Papers in Refereed Conference Proceedings:**

Ek, A., & Bernardy, J. P. (2020), “Composing Byte-Pair Encodings for Morphological

Sequence Classification”, *Proceedings of the Fourth Workshop on Universal Dependencies*

*(UDW 2020)*, pp. 76-86.

Ek, A., & Bernardy, J. P. (2020), “How Much of Enhanced UD is Contained in UD?”,

*Proceedings of the 16th International Conference on Parsing Technologies and the IWPT*

*2020*, Shared Task on Parsing into Enhanced Universal Dependencies, pp. 221-226.

Ek, A., Bernardy, J. P., & Chatzikyriakidis, S. (2020), “How does Punctuation Affect Neural

Models in Natural Language Inference”, *Proceedings of the Probability and Meaning*

*Conference (PaM 2020)*, ACL Anthology, pp. 109-116,

**Invited Talks and Seminars:**

Enhanced Dependency Parsing, CLASP Seminar, March 11, 2020.

Composing Byte-Pair Encodings for Morphological Sequence Classification, CLASP Seminar, November 20, 2020

**PhD Thesis Work:**

Completed three more papers for my thesis (4 papers currently). 2-3 more to go, which should be done by the end of 2021/spring 2022. Half of my course credits will soon be completed.

**Tewodros Gebreselassie**

**Papers in Refereed Conference Proceedings:**

Abate, Solomon Teferra; Tachbelie, Martha Yifiru; Melese, Michael; Abera,

Hafte; Gebreselassie, Tewodros; Mulugeta, Wondwossen; Assabie, Yaregal;

Beyene, Million Meshesha; Atinafu, Solomon; Seyoum, Binyam Ephrem (2020),

“Large Vocabulary Read Speech Corpora for Four Ethiopian Languages: Amharic,

Tigrigna, Oromo, and Wolaytta”, *Proceedings of the The Fourth Widening Natural*

*Language Processing Workshop*, pp. 13-17, ACL Anthology.

**Workshop Organising:**

Co-organiser of the Workshop on Practical ML for Developing Countries: Learning under Limited/Low Resource Scenarios, ICLR 2020, The Eighth International Conference on Learning Representations

Co-organiser of the RESOURCEFUL-2020 (RESOURCEs and Representations for Under-Resourced Languages and Domains) workshop at SLTC

**PhD Thesis Work:**

Tewodros collected data in Wolaytta, an Ethiopian language, in particular a list of words and a list of sentences that can be used for computational work on the semantic similarity task.

Over the summer he returned to Ethiopia to be with his family but his stay was prolonged and is still ongoing because of the difficult political situation that emerged since then coupled with the pandemic. It has been hard for him to work online because of the unstable internet connection in the country. Therefore, he primarily used this time for onsite data collection, while he is planning to do more NLP focused work once he returns to Sweden in the coming months.

**Julian Grove**

**Seminar:**

Algebraic Effects in Montague Semantics, CLASP Seminar, University of Gothenburg, October 28, 202

**Christine Howes**

**Journal Article:**

Gregoromichelaki, E., Mills, G. J., Howes, C., Eshghi, A., Chatzikyriakidis, S., Purver, M. Kempson R., Cann, R. and Healey, P. (2020), “Completability vs (In)completeness”, *Acta Linguistica Hafniensia*.

**Book Chapters:**

Gregoromichelaki, E., Kempson, R. & Howes, C (2020), “Actionism in Syntax and

Semantics”, in Howes, C., Dobnik, S. & Breitholtz, E. (editors), *Dialogue and Perception –*

*Extended Papers from DaP2018*, Gothenburg : GUPEA.

Mazzocconi, C., Jin, G., Maraev, V., Howes, C., Ginzburg, J. & Scott, S (2020). Laughables

and laughter perception: Preliminary investigations. In Howes, C., Dobnik, S. & Breitholtz,

E. (editors), *Dialogue and Perception - Extended papers from DaP2018*, Gothenburg :

GUPEA.

**Papers in Refereed Conference Proceedings:**

Breitholtz, E. & Howes, C (2020), “Communicable Reasons: How Children Learn Topoi

through Dialogue, *Proceedings of the 24th Workshop on the Semantics and Pragmatics of*

*Dialogue*, Waltham, MA : SEMDIAL.

Breitholtz, E. & Howes, C (2020), “Theirs Not to Reason Why: Dialogical Reasoning for

Conversational Artificial Agents*, Do Robots Talk? Philosophical Implications of*

*Describing Human-Machine Communication (DoRoTa)*.

Dobnik, S., Kelleher, J. & Howes, C (2020), “Local Alignment of Frame of Reference

Assignment in English and Swedish Dialogue”, Šķilters, J., Newcombe, N. S. & Uttal, D.

(editors), *Spatial Cognition {XII}: Papers from 12th International Conference, Spatial*

*Cognition 2020/21*, Riga, Latvia.

Gregoromichelaki, E., Chatzikyriakidis, S., Eshghi, A., Hough, J., Howes, C., Kempson, R.,

Kiaer, J., Purver, M., Sadrzadeh, M. and White, G. 2020, “Affordance Competition in

Dialogue: The Case of Syntactic Universals,” P*roceedings of SemDial* (WatchDial2020).

Larsson, S., Howes, C. & Bondarenko, A (2020), “Could You Spell that Again Please?

Towards a Formal Model of Grounding in Directory Enquiries”, *First AISB Symposium on*

*Conversational AI (SoCAI)*.

Maraev, V., Breitholtz, E., and Howes, C. (2020a), "How Do You Make an AI Get the Joke?

Here’s What I Found on the Web", *First AISB Symposium on Conversational AI (SoCAI)*.

Maraev, V., Mazzocconi, C., Mills, G., and Howes, C. (2020b), "LOL What?: *Empirical*

*Study of Laughter in Chat Based Dialogues, Proceedings of Laughter Workshop 2020*.

Mazzocconi, C., Jin, G., Maraev, V., Howes, C., Ginzburg, J., and Scott, S. (2020),

“Laughables and Laughter Perception: Preliminary Investigations” in Howes, C., Dobnik,

S., and Breitholtz, E. (editors), *Dialogue and Perception - Extended papers from DaP2018*,

volume 2 of CLASP papers in computational linguistics, pages 72–83. GUPEA,

Gothenburg.

Somashekarappa, V., Howes, C. & Sayeed, A (2020), “An Annotation Approach for Social

and Referential Gaze in Dialogue”, *Proceedings of the Twelfth International Conference*

*on Language Resources and Evaluation (LREC)*.

**Edited Volumes:**

Howes, C., Dobnik, S. & Breitholtz, E., editors (2020), *Dialogue and Perception - Extended*

*Papers from DaP2018*, CLASP papers in computational linguistics, Gothenburg, GUPEA.

Howes, C., Chatzikyriakidis, S., Ek, A. & Somashekarappa, V., editors (2020), *Proceedings*

*of the Probability and Meaning Conference (PaM 2020)*, Gothenburg, Association for

Computational Linguistics.

**Conference Organising:**

CLASP Probability and Meaning conference (PaM), held virtually in October 2020.

**Nikolai Ilinykh**

**Papers in Refereed Conference Proceedings:**

Ilinykh Nikolai, & Dobnik Simon (2020), “When an Image Tells a Story: The Role of Visual

and Semantic Information for Generating Paragraph Descriptions”, *Proceedings of the*

*13th International Conference on Natural Language Generation*, Association for

Computational Linguistics, pp. 338–348

Moussallem Diego, Kaur Paramjot, Ferreira Thiago, Lee Chris van der, Shimorina Anastasia,

Conrads Felix, Röder Michael, Speck René, Gardent Claire, Mille Simon, Ilinykh Nikolai,

& Ngonga Ngomo Axel-Cyrille (2020), “A General Benchmarking Framework for Text

Generation”, *Proceedings of the 3rd International Workshop on Natural Language*

*Generation from the Semantic Web (WebNLG+)*, Association for Computational

Linguistics, pp. 27–33.

Castro Ferreira Thiago, Gardent Claire, Ilinykh Nikolai, Lee Chris van der, Mille Simon,

Moussallem Diego, & Shimorina Anastasia. (2020), “The 2020 Bilingual, Bi-Directional

WebNLG+ Shared Task: Overview and Evaluation Results (WebNLG+ 2020)”,

*Proceedings of the 3rd International Workshop on Natural Language Generation from the*

*Semantic Web (WebNLG+),* Association for Computational Linguistics, pp. 55–76.

**Conference Organising:**

Organizer and program committee member for the 3rd Workshop on Natural Language Generation from Semantic Web (WebNLG+), 15-18 December, Dublin, Ireland (virtual).

Member of the organizing committee for the WebNLG+ Challenge 2020.

**PhD Thesis Work:**

In 2020 I built a practical and theoretical foundation for my dissertation: what type of data, models and tasks I want to focus on. I have explored several ways to represent input data for encoder-decoder modelling with papers on text generation from visual and meaning representations. I have also attended the course on transformers, and currently working on the problems of explainability and interpretability of such architectures for image captioning. Finally, given my previous research experience on more interactive tasks (Visual Dialogue), I started looking into how I can extend my interpretability-based analysis of the models to a more complex human-robot interaction setting.

**Aram Karimi**

Maryam Rajestari, Simon Dobnik, Robin Cooper and Aram Karimi (2020), "Very

Necessary: the Meaning of Non-Gradable Modal Adjectives in Discourse Contexts",

*Proceedings of the Eighth Swedish Language Technology Conference (SLTC)*, 25–27

November, Gothenburg.

**Research Programming Support**

Assisted in Asad Sayeed’s project to predict proto role labeling and run it on some ukwac examples. This work was finished in April, and then provided support for docker installation.

Worked with Bill Noble to construct the CLASP external website

Working with Jean-Philippe Bernardy on support and administration of the CLASP Lark CPU + GPU server

Cooperating with Simon Dobnik on the development of web chat tools

Supporting Simon Dobnik’s development of a scientific recommender system using BERT

**Shalom Lappin**

**Article in Refereed Journal:**

Jey Han Lau, Carlos Armendariz, Shalom Lappin, Matthew Purver, and Chang Shu (2020),

"How Furiously Can Colorless Green Ideas Sleep? Sentence Acceptability in

Context", *Transactions of the Association for Computational Linguistics*8, pp.  296—310

(presented at the 58th Annual Conference of the Association for Computational Linguistics.

July, 2020).

**Invited Course:**

Deep Learning and the Nature of Linguistic Representation, WeSSLLI, Brandeis, July 11-17, 2020.

This course provided the basis for a monograph, Shalom Lappin, *Deep Learning and* *Linguistic Representation*, that Taylor and Francis is publishing in May, 2021.

**Invited Talks and Seminars:**

Towards a Bayesian Semantics for Natural Language, Programming Principles, Logic, and Verification Group Research Seminar, Computer Science, UCL, January 23, 2020

Predicting Human Judgments of Sentence Acceptability in Context, Language Technology Lab Seminar, University of Cambridge, February 6, 2020

Predicting Human Judgments of Sentence Acceptability in Context, Rutherford Appleton Laboratory Machine Learning Seminar, February 27, 2020

Deep Learning and Sentences in Context, Integrating Generic and Contextual Knowledge Symposium (GeCKo), May 18, 2020, Barcelona

**Staffan Larsson**

**Papers in Refereed Conference Proceedings:**

Jenny Myrendal & Staffan Larsson (2020), “Ordets Makt i Klimatdebatten”, *Forum för*

*Bioekonomi*, 1/2020, pp. 50-51.

Larsson, S., Howes, C. & Bondarenko, A (2020), “Could You Spell that Again Please?

Towards a Formal Model of Grounding in Directory Enquiries”, *First AISB Symposium on*

*Conversational AI (SoCAI)*.

Staffan Larsson (2020), “Extensions Are Indeterminate if Intensions Are Classifiers”,

*Proceedings of the 24th Workshop on the Semantics and Pragmatics of Dialogue*, held

virtually at Brandeis University, Waltham, Massachusetts July 18-19 2020 / Sophia

Malamud, James Pustejovsky & Jonathan Ginzburg (eds.).

Staffan Larsson (2020), “Discrete and Probabilistic Classifier-based Semantics”, *Proceedings*

*of the Probability and Meaning Conference (PaM 2020)*, Gothenburg and online, 14–15

October 2020 / Stergios Chatzikyriakidis, Christine Howes (Editors).

**Invited Talks and Seminars:**

Jenny Myrendal & Staffan Larsson (2020), Rörliga Mål - Ordbetydelseförhandling i Dialog, Vetenskapsfestivalen Göteborg 2020.

**Invited Course:**

(with Robin Cooper), Modelling Linguistic Communication Using Types, WeSSLLI, Brandeis, 11th-17th July, 2020.  
  
**Workshop Organising:**

Local co-organiser of the Eighth Swedish Language Technology Conference (SLTC), <https://spraakbanken.gu.se/en/sltc2020>

**Vlad Maraev**

**Papers in Refereed Conference Proceedings:**

Maraev, V., Breitholtz, E., and Howes, C. (2020a), "How Do You Make an AI Get the Joke?

Here’s What I Found on the Web", *First AISB Symposium on Conversational AI (SoCAI)*.

Maraev, V., Mazzocconi, C., Mills, G., and Howes, C. (2020b), "LOL What?: *Empirical*

*Study of Laughter in Chat Based Dialogues, Proceedings of Laughter Workshop 2020*.

Mazzocconi, C., Jin, G., Maraev, V., Howes, C., Ginzburg, J., and Scott, S. (2020),

“Laughables and Laughter Perception: Preliminary Investigations” in Howes, C., Dobnik,

S., and Breitholtz, E. (editors), *Dialogue and Perception - Extended papers from DaP2018*,

volume 2 of CLASP papers in computational linguistics, pages 72–83. GUPEA,

Gothenburg.

**Seminar Talks:**

(with Bill Noble) Tickle me BERT: Using Laughter to Probe Pre-Trained Encoders for Dialogue Act Recognition, CLASP seminar, February 26, 2020

**PhD Thesis Work:**

In 2020 I worked on the text of the monograph’s manuscript and submitted two journal articles, one of those is under review and another is accepted for publication in 2021. The tentative deadline for the final full draft is set for the end of June 2021.

**Bill Noble**

**Papers in Refereed Conference Proceedings:**

Noble, Bill, Ellen Breitholtz and Robin Cooper (2020), Personae under Uncertainty: The case

of Topoi, in Stergios Chatzikyriakidis, Christine Howes (Editors), *Proceedings of the*

*Probability and Meaning Conference (PaM 2020)*, SiGSEM, ACL Anthology.

**Invited Talks and Seminars:**

The Effect of Laughter on Dialogue Act Recognitionm CLASP Seminar, 26 Feb 2020

Conditional Language Models for Linguistic Variation and Change, 2nd Workshop on Computational Detection of Language Change, 25 Nov 2020

**PhD Thesis Work:**

Dissertation progress has focused on two main projects.

(i) Semantic shift in social networks: We use distributional semantics to test hypotheses about the relationship between semantic shift and social network structure. Began work in February 2020 and submitted to EACL in October. Currently revising for resubmission to \*-SEM or another ACL workshop.

(ii) Word meaning negotiation: Began work in 2019, but re-started in Fall of 2020. Developed graph model of word meaning negotiations with nodes as "semantic anchors" and edges as proposed relations between them. Developed an annotation scheme with MLT student Catherine Viloria, as part of her research internship course.

**Additional Research Activity:**

Distributional discourse representation: This work is in connection to the CLASP seminar Vlad Maraev and I gave on laughter and dialogue act recognition. We experimented with extending a BERT sentence (utterance) encoder with a RNN model to represent the dialogue state, and inferring dialogue act from the dialogue state representation. We are now working on a workshop submission based on that.

Social meaning and topoi: Work with Ellen Breitholtz and Robin Cooper, related to our PaM paper. We are extending that work for a journal article submission, and I will also include concepts from that work, especially the relationship with rational speech acts, in the book chapter with Vlad.

**Kathrein Qwaider**

**Papers in Refereed Conference Proceedings:**

Qwaider, C., Saad, M., Chatzikyriakidis, S. and Dobnik S. and Johansson. R. (2020), “An

Arabic Tweets Sentiment Analysis Dataset (ATSAD) Using Distant Supervision and Self

Training”, P*roceedings of the 4th Workshop on Open-Source Arabic Corpora and*

*Processing Tools*, LREC2020.

**Seminar:**

Kathrein Qwaider, Deep Learning for Arabic Computational Linguistics (Sentiment Analysis as case study), CLASP Seminar, April 1, 2020

**PhD Thesis Work:**

Kathrein is writing a compilation thesis. 5 papers so far have been published, one is to be resubmitted, and we are hoping for another one. She has finished all her course work. Chatrine has been on parental leave since September 2020.

**Sylvie Saget**

**PhD Thesis Work:**

Project Title: Language as a tool: A Rational and Cognitive Approach

Starting from a study of Common Ground, dissertation progress has focussed on:

(1) Specification of a rational and cognitive agent framework based on the distinction two types of information (facts and information related to what to do or use/how to behave) and the specification of corresponding information containers (doxastic mental attitude and memory-box). A multidisciplinary approach is used.

(2) Dialog Modeling: How does our rational and cognitive framework contribute to study and explain elements such as alignments or conventions, reasoning processes such as how people understand each other's or argumentation?

(3) Conversational AI: How our framework may be used to enhance Conversational AI/Social Robotics/Agents systems?

A detailed outline of PhD thesis has been developed.

**Additional Research Activity:**

Following Bec Johnson(University of Melbourne)'s invitation, Sylvie joined the "PhD students in AI Ethics" group (https://phdaiethics.com/).

**Asad Sayeed**

**Journal Article:**

Kågebäck, M., Carlsson, E., Dubhashi, D., & Sayeed, A. (2020), “A Reinforcement-Learning

Approach to Efficient Communication, *Plos One*, 15(7), e0234894.

**Papers in Refereed Conference Proceedings:**

Somashekarappa, V., Howes, C., & Sayeed, A. (2020), “An Annotation Approach for Social

and Referential Gaze in Dialogue”, *Proceedings of The 12th Language Resources and*

*Evaluation Conference*, pp. 759-765.

Loáiciga, S., Hardmeier, C., & Sayeed, A. (2020), “Exploiting Cross-Lingual Hints to

Discover Event Pronouns”, *Proceedings of The 12th Language Resources and Evaluation*

*Conference*, pp. 99-103.

Olsen, I. R., Pedersen, B. S., & Sayeed, A. (2020), “Building Sense Representations in Danish

by Combining Word Embeddings with Lexical Resources”, *Proceedings of the 2020*

*Globalex Workshop on Linked Lexicography*, pp. 45-52.

Hong, X., Shetty, R., Sayeed, A., Mehra, K., Demberg, V., & Schiele, B. (2020), “Diverse

and Relevant Visual Storytelling with Scene Graph Embeddings”, *Proceedings of the 24th*

*Conference on Computational Natural Language Learning*. pp. 420-430.

**Invited Talks:**

Sayeed, A. GRIPES: Gothenburg research initiative for politically emergent systems. (2020, September).  WASP-HS online meeting.

Sayeed, A., Adversarial vagueness: manipulative language and autonomous agents. (2020, October).  AI and Humanities public lecture series, University of Gothenburg.

Sayeed, A. Dogwhistles, manipulative language, and autonomous agents. (2020, December).  Invited talk at Swedish School of Library and Information Science, University of Borås.

Asad was also promoted to Senior Lecturer, and has a simultaneous application for docent status still pending, with a positive evaluation already returned

**Vidya Somashekarappa**

**Papers in Refereed Conference Proceedings:**

Vidya Somashekarappa, Christine Howes, & Asad Sayeed. (2020). “An Annotation Approach

for Social and Referential Gaze in Dialogue”, *Proceedings of* *the 12th Language Resources*

*and Evaluation Conference* (pp. 759-765), Marseille, France, May 2020, The Association

for Computational Linguistics.

**Edited Volumes:**

Christine Howes, Stergios Chatzikyriakidis, Adam Ek & Vidya Somashekarappa (2020),

editors, *Proceedings of the Conference on Probability and Meaning (PaM).* Gothenburg,

Sweden and online, 14–15 October 2020. The Association for Computational Linguistics.

**Conference Organising:**

Local organiser and programme co-chair (with Christine Howes, Stergios Chatzikyriakidis and Adam Ek) of the Conference on Probability and Meaning (PaM). Gothenburg, Sweden and online, 14–15 October 2020.

**Other Research Activity:**

Deputy PhD representative of the steering committee of the Centre for Digital Humanities/Centrum för digital humanioras.

Member of the Humanities Doctoral Council/Humanistiska doktorandrådet (HDR)

Participated in the Annual autumn workshop of Nationella språkbanken, October 2020

Participated in the 8th Swedish Language Technology Conference (SLTC), 25-27 October 2020

**Travel Grant:**

Adlerbertska Stipendier travel grant (20 000 SEK)

**PhD Thesis Work:**

I'll be writing a monograph thesis. So far, we have a video corpus annotated for around 50 minutes of speech, gaze and 14 minutes of laughter. A detailed qualitative and quantitative analysis of the same has been done. I'm now working on automating these annotations along with modelling the attributes on Furhat (avatar) which is going to be the task for this year.