Introduction

This volume contains the papers presented at the CLASP midterm conference, Probability and Meaning (PaM2020) at the Department of Philosophy, Linguistics and Theory of Science (FLoV), University of Gothenburg, held on October 14-15th, 2020.

PaM brings together researchers interested in computationally relevant probabilistic approaches to natural language semantics and includes symbolic, machine learning and experimental approaches to this task, as well as hybrid models.

Papers were invited on topics in these and closely related areas, including (but not limited to) probabilistic approaches developed within a computational framework, the semantics of natural language for written, spoken, or multimodal communication, probabilistic type theoretic approaches to meaning, multimodal and grounded approaches to computing meaning, dialogue modelling and linguistic interaction, deep learning approaches and probability, syntax-semantics interface, alternative approaches to compositional semantics, inference systems for computational semantics, recognising textual entailment, semantic learning, computational aspects of lexical semantics, semantics and ontologies, semantic aspects of language generation, semantics-pragmatics interface.

This conference aims to initiate a genuine discussion between these related areas and to examine different approaches from computational, linguistic and psychological perspectives and how these can inform each other. It features 4 invited talks by leading researchers in these areas, and 18 peer-reviewed papers, 11 presented as long talks and 7 presented as posters.

We would like to thank all our contributors and programme committee members, with special thanks to CLASP for organising the virtual conference and our sponsors SIGSEM http://sigsem.org, the ACL special interest group on semantics, and the Swedish Research Council (Vetanskaprådet) for funding CLASP.

Christine Howes, Stergios Chatzikyriakidis, Adam Ek and Vidya Somashekarappa

Gothenburg

October 2020
Program Committee:

Rodrigo Agerri  
Anja Belz  
Emily M. Bender  
Raffaella Bernardi  
Jean-Philippe Bernardy  
Johan Bos  
Ellen Breitholtz  
Harry Bunt  
Aljoscha Burchardt  
Nicoletta Calzolari  
Rui Chaves  
Alexander Clark  
Stephen Clark  
Ariel Cohen  
Robin Cooper  
Philippe de Groote  
Leon Derczynski  
Markus Egg  
Adam Ek  
Katrin Erk  
Arash Eshghi  
Jonathan Ginzburg  
Julian Hough  
Elisabetta Jezek  
Richard Johansson  
John Kelleher  
Ralf Klabunde  
Emiel Krahmer  
Steffan Larsson  
Vladislav Maraev  
Paul McKevitt  
Louise McNally  
Marie-Francine Moens  
Shashi Narayan  
Joakim Nivre  
Bill Noble  
Denis Paperno  
Anselmo Peñas  
Manfred Pinkal  
Massimo Poesio  
Violaine Prince  
Stephen Pulman  
Matthew Purver  
James Pustejovsky  

University of the Basque Country  
University of Brighton  
University of Washington  
University of Trento  
University of Gothenburg  
University of Groningen  
University of Gothenburg  
Tilburg University  
German Research Center for Artificial Intelligence (DFKI GmbH)  
National Research Council of Italy  
University of Buffalo  
King's College London  
Queen Mary University of London  
Ben-Gurion University of the Negev  
University of Gothenburg  
Inria  
IT University of Copenhagen  
Humboldt University  
University of Gothenburg  
University of Texas at Austin  
Heriot-Watt University  
Université Paris-Diderot  
Queen Mary University of London  
University of Pavia  
Chalmers University of Technology  
Technological University Dublin  
Ruhr-Universität Bochum  
Tilburg University  
University of Gothenburg  
University of Gothenburg  
University of Gothenburg  
Ulster University  
Universitat Pompeu Fabra  
KU Leuven  
Google  
Uppsala University  
University of Gothenburg  
Loria, CNRS  
Universidad Nacional de Educación a Distancia  
Saarland University  
Queen Mary University  
University of Montpellier  
Oxford University  
Queen Mary University of London  
Brandeis University
Program Committee (continued):

Christian Retoré University of Montpellier
German Rigau University of the Basque Country
Hannah Rohde University of Edinburgh
Mehrooosh Sadrzadeh University College London
Asad Sayeed University of Gothenburg
David Schlangen University of Potsdam
Sabine Schulte im Walde University of Stuttgart
Vidya Somashekarappa University of Gothenburg
Tim Van de Cruys University of Toulouse
Eva Maria Vecchi University of Cambridge
Carl Vogel Trinity College Dublin

Invited Speakers:

Heather Burnett, Laboratoire de Linguistique Formelle, Université de Paris
Stephen Clark, Department of Computer Science and Technology, University of Cambridge
Katrin Erk, Linguistics Department, University of Texas at Austin
Noah Goodman, Departments of Computer Science and Psychology, Stanford, CA
**Invited talk 1: Heather Burnett**

**Social Signaling and Reasoning under Uncertainty: French “Écriture Inclusive”**

Gender inclusive writing (“écriture inclusive” EI) has long been the topic of public debates in France. Examples of EI for the word “students” are shown in (1).

(1) a. étudiant-e-s (point médian)
   b. étudiant.e.s (period)
   c. étudiants et étudiantes (repetition)
   d. étudiant(e)s (parentheses)
   e. étudiant-e-s (dash)
   f. étudiantEs (capital)
   g. étudiant/e/s (slash)
   h. étudiant- -e- -s (double dash)

These debates have amplified since the Macron government prohibited the use of the point médian (1a) in official documents in 2017 (Abbou et al. 2018). In addition to being a point of disagreement between feminists and anti-feminists, EI is also controversial among feminists: it has many variants (1), who often disagree on which variant should be used (Abbou 2017). In this talk, I argue that the source of many of these disagreements lies in the fact that French écriture inclusive has developed into a rich social signalling system: based on a quantitative study of EI in Parisian university brochures (joint work with Céline Pozniak (Burnett & Pozniak 2020)), I argue that writers use or avoid EI in part in order to communicate aspects of their political orientations. We show that these aspects involve writers’ perspectives on gender, but also stances on issues unrelated to gender, such as (anti)institutionalism and support for the Macron government. I then outline a research programme for studying this signalling system from a formal perspective: following Burnett (2019), I show how we can use probabilistic pragmatics to analyze EI’s contribution to writers’ political identity construction and the consequences that this has for its use as a tool for promoting gender equality and social change.

**Invited talk 2: Katrin Erk**

**How to marry a star: Probabilistic constraints for meaning in context**

Context has a large influence on word meaning; not only local context, like in the combination of a predicate and its argument, but also global topical context. In computational models, this is routinely factored in, but the question of how to integrate different context influences is still open for theoretical accounts of sentence meaning. We start from Fillmore’s "semantics of understanding", where he argues that listeners expand on the "blueprint" that is the original utterance, imagining the utterance situation by using all their knowledge about words and the world. We formalize this idea as a two-tier "situation description system" that integrates referential and conceptual representations of meaning.

A situation description system is a Bayesian generative model that takes utterance understanding to be the mental process of probabilistically describing one or more situations that would make a speaker’s utterance logically true, from the point of view of the listener.
Invited talk 3: Stephen Clark

Grounded Language Learning in Virtual Environments

Natural Language Processing is currently dominated by the application of text-based language models such as BERT and GPT. One feature of these models is that they rely entirely on the statistics of text, without making any connection to the world, which raises the interesting question of whether such models could ever properly “understand” the language. One way in which these models can be grounded is to connect them to images or videos, for example by conditioning the language models on visual input and using them for captioning. In this talk I extend the grounding idea to a simulated virtual world: an environment which an agent can perceive, explore and interact with. More specifically, a neural-network-based agent is trained – using distributed deep reinforcement learning – to associate words and phrases with things that it learns to see and do in the virtual world. The world is 3D, built in Unity, and contains recognisable objects, including some from the ShapeNet repository of assets.

One of the difficulties in training such networks is that they have a tendency to overfit to their training data, so first we’ll demonstrate how the interactive, first-person perspective of an agent provides it with a particular inductive bias that helps it to generalize to out-of-distribution settings. Another difficulty is that training the agents typically requires a huge number of training examples, so we’ll show how meta-learning can be used to teach the agents to bind words to objects in a one-shot setting. Moreover, the agent is able to combine its knowledge of words obtained one-shot with its stable knowledge of word meanings learned over many episodes, providing a form of grounded language learning which is both “fast and slow”.

Joint work with Felix Hill.

Invited talk 4: Noah Goodman

Reference, Inference, and Learning

A key function of human language is reference to objects and situations. Referential language grounds in stable semantic conventions, but flexibly depends on context. In this talk I will explore the computational mechanisms of referential language in the setting of language games. I will argue that many patterns of behavioral data can be explained by a combination of hierarchical learning for semantics – realized with the tools of deep neural networks – and recursive social reasoning for pragmatics – realized in the Bayesian rational speech acts (RSA) framework. I will consider phenomena of redundancy in reference, grounding semantics in vision, and adaptation under repeated interaction. Finally I will address a key puzzle for RSA (and other neo-Gricean theories): how can production be so quick and effortless if it depends on complex recursive reasoning?
# Table of Contents

'Practical’, if that's the word  
Eimear Maguire ................................................................. 1

Personae under uncertainty: The case of topoi  
Bill Noble, Ellen Breitholtz and Robin Cooper .......................... 8

Dogwhistles as Identity-based interpretative variation  
Quentin Dénigot and Heather Burnett ................................. 17

Conditional answers and the role of probabilistic epistemic representations  
Jos Tellings ........................................................................... 26

Linguistic interpretation as inference under argument system uncertainty: the case of epistemic must  
Brandon Waldon ................................................................. 34

Linguists Who Use Probabilistic Models Love Them: Quantification in Functional Distributional Semantics  
Guy Emerson ......................................................................... 41

Fast visual grounding in interaction: bringing few-shot learning with neural networks to an interactive robot  
José Miguel Cano Santín, Simon Dobnik and Mehdi Ghanimifard ................................. 53

Discrete and Probabilistic Classifier-based Semantics  
Staffan Larsson ....................................................................... 62

Social Meaning in Repeated Interactions  
Elin McCready and Robert Henderson .................................. 69

Towards functional, agent-based models of dogwhistle communication  
Robert Henderson and Elin McCready .................................. 73

Stochastic Frames  
Annika Schuster, Corina Stroessner, Peter Sutton and Henk Zeevat ................................. 78

A toy distributional model for fuzzy generalised quantifiers  
Mehrnoosh Sadrzadeh and Gijs Wijnholds ............................ 86

Generating Lexical Representations of Frames using Lexical Substitution  
Saba Anwar, Artem Shelmanov, Alexander Panchenko and Chris Biemann .................... 95

Informativity in Image Captions vs. Referring Expressions  
Elizabeth Coppock, Danielle Dionne, Nathaniel Graham, Elias Ganem, Shijie Zhao, Shawn Lin, Wenxing Liu and Derry Wijaya .................................................. 104

How does Punctuation Affect Neural Models in Natural Language Inference  
Adam Ek, Jean-Philippe Bernardy and Stergios Chatzikyriakidis .................................. 109

Building a Swedish Question-Answering Model  
Hannes von Essen and Daniel Hesslow .................................. 117

Word Sense Distance in Human Similarity Judgements and Contextualised Word Embeddings  
Janosch Haber and Massimo Poesio ...................................... 128
Conference Program

October 14th

Session 1

'Practical', if that's the word
Eimear Maguire

*Personae under uncertainty: The case of topoi*
Bill Noble, Ellen Breitholz and Robin Cooper

*Dogwhistles as Identity-based interpretative variation*
Quentin Dénigot and Heather Burnett

Session 2

*Conditional answers and the role of probabilistic epistemic representations*
Jos Tellings

*Linguistic interpretation as inference under argument system uncertainty: the case of epistemic must*
Brandon Waldon

Session 3

*Linguists Who Use Probabilistic Models Love Them: Quantification in Functional Distributional Semantics*
Guy Emerson

*Fast visual grounding in interaction: bringing few-shot learning with neural networks to an interactive robot*
José Miguel Cano Santín, Simon Dobnik and Mehdi Ghanimifard
October 15th

Poster session

Discrete and Probabilistic Classifier-based Semantics
Staffan Larsson

Social Meaning in Repeated Interactions
Elin McCready and Robert Henderson

Towards functional, agent-based models of dogwhistle communication
Robert Henderson and Elin McCready

Stochastic Frames
Annika Schuster, Corina Stroessner, Peter Sutton and Henk Zeevat

A toy distributional model for fuzzy generalised quantifiers
Mehrnoosh Sadrzadeh and Gijs Wijnholds

Generating Lexical Representations of Frames using Lexical Substitution
Saba Anwar, Artem Shelmanov, Alexander Panchenko and Chris Biemann

Informativity in Image Captions vs. Referring Expressions
Elizabeth Coppock, Danielle Dionne, Nathanial Graham, Elias Ganem, Shijie Zhao,
Shawn Lin, Wenxing Liu and Derry Wijaya
Session 4

*How does Punctuation Affect Neural Models in Natural Language Inference*
Adam Ek, Jean-Philippe Bernardy and Stergios Chatzikyriakidis

*Building a Swedish Question-Answering Model*
Hannes von Essen and Daniel Hesslow

Session 5

*Word Sense Distance in Human Similarity Judgements and Contextualized Word Embeddings*
Janosch Haber and Massimo Poesio

*Short-term Semantic Shifts and their Relation to Frequency Change*
Anna Marakasova and Julia Neidhardt