

# Discrete and Probabilistic Classifier-based Semantics

Staffan Larsson

Centre for Linguistic Theory and Studies in Probability (CLASP)

Department of Philosophy, Linguistics and Theory of Science

University of Gothenburg, Sweden

sl@ling.gu.se

## Abstract

We present a formal semantics (a version of Type Theory with Records) which places classifiers of perceptual information at the core of semantics. Using this framework, we present an account of the interpretation and classification of utterances referring to perceptually available situations (such as visual scenes). The account improves on previous work by clarifying the role of classifiers in a hybrid semantics combining statistical/neural classifiers with logical/inferential aspects of meaning. The account covers both discrete and probabilistic classification, thereby enabling learning, vagueness and other non-discrete linguistic phenomena.

## 1 Introduction

Marconi (1997) distinguishes inferential and referential meaning. *Inferential* word meanings enable inferences from uses of the word. Such meanings are sometimes referred to as “high level” or “symbolic”, and are typically modelled in formal semantics. *Referential* meaning, on the other hand, allows speakers to identify objects and situations referred to. Referential meaning is sometimes referred to as “low-level” or “subsymbolic”. Our working hypothesis is that referential meaning can be modelled using classifiers that output formal representations (Larsson, 2011, 2015), thus connecting “high level” formal representations to “low level” perceptual information. This is a way of addressing the *symbol grounding problem* put forward by (Harnad, 1990) in a way that is compatible with formal semantics.

We also want a framework where meanings can be learned from interactions, and where dialogue participants can coordinate on meanings (Larsson and Myrendal, 2017). To enable this, intensions need to be represented independently of extensions as structured objects which can be modi-

fied (updated), and include classifiers of perceptual data.

In formal semantics in the Montague tradition (Montague, 1974), the meaning of a word such as “dog” is taken to be its extension, i.e. the set of all dogs in the world (or in a possible world). This type of semantic theory does not represent intensions independently of extensions, which makes it less well suited for modelling aspects of referential meaning using classifiers. For example, modelling a classifier extensionally (as, say, a set of ordered pairs of inputs and outputs) seems to require some external classifier procedure to produce these sets. Taken as a model of natural language meaning, this suggests an counter-intuitive and unrealistic account of how humans encounter new situations and classify them. Furthermore, such a theory would exclude classification and classification learning (which we take to be part of the acquisition of word meanings) from semantic theory proper, when we in fact believe that it is central to semantics. For these reasons and others, traditional Montagovian semantics does not seem to us to be a satisfactory framework for classifier-based semantics. However, we do believe that it is crucial that the accumulated insights from work in formal semantics over the last 50 decades are integrated with the ideas put forward in this paper.

## 2 Background

We are developing a formal *judgement-based semantics* where notions such as perception, classification, judgement, learning and dialogue coordination play a central role (Cooper, 2005; Larsson and Cooper, 2009; Larsson, 2011; Dobnik et al., 2011; Cooper, 2012; Dobnik and Cooper, 2013; Cooper et al., 2015). A key idea introduced in Larsson (2011) and Larsson (2015) is the modelling of referential meanings as classifiers of real-

valued (perceptual) data, and training these classifiers in interaction with the world and other agents.

There is a growing body of work in computational and formal semantics which is in line with the approach taken here (Kennington and Schlangen, 2015; Andreas et al., 2016; Schlangen et al., 2016; Ghanimifard and Dobnik, 2017; Shore and Skantze, 2018). We propose a way of connecting this line of work to formal semantics, to enable combining it with the successes of formal semantics (compositionality, quantification, etc.).

Using a Type Theory with Records (Cooper et al., 2014), Larsson (2015) presents a formal semantics for perception, using classifiers to model the relation between perception and linguistic utterances. This paper substantially improves on the formal machinery used in Larsson (2015) and incorporates insights from the implemented version of TTR (Cooper, 2019) as well as related work on visual question answering (Utescher, 2019).

### 3 TTR: A brief introduction

We will be formulating our account in a Type Theory with Records (TTR). We can here only give a brief and partial introduction to TTR; see also Cooper (2005) and Cooper (2012). To begin with,  $s : T$  is a judgment that some  $s$  is of type  $T$ . To make explicit who is making this judgment, the of-type relation may be subscripted with an agent  $A$ , as in  ${}_A T$ . One *basic type* in TTR is  $\text{Ind}$ , the type of an individual; another basic type is  $\mathbb{R}$ , the type of real numbers. Given that  $T_1$  and  $T_2$  are types,  $T_1 \rightarrow T_2$  is a *functional type* whose domain is objects of type  $T_1$  and whose range is objects of type  $T_2$ .

Next, we introduce *records* and *record types*. If  $a_1 : T_1, a_2 : T_2(a_1), \dots, a_n : T_n(a_1, a_2, \dots, a_{n-1})$ , where  $T(a_1, \dots, a_n)$  represents a type  $T$  which depends on the objects  $a_1, \dots, a_n$ , the record to the left in Figure 1 is of the record type to the right.

In Figure 1,  $\ell_1, \dots, \ell_n$  are *labels* which can be used elsewhere to refer to the values associated with them. A sample record and record type is shown in Figure 2.

Types constructed with predicates may be *dependent*. This is represented by the fact that arguments to the predicate may be represented by labels used on the left of the ‘:’ elsewhere in the record type. In Figure 2, the type of  $c_{\text{man}}$  is dependent on  $\text{ref}$  (as is  $c_{\text{run}}$ ).

If  $r$  is a record and  $\ell$  is a label in  $r$ , we can use a *path*  $r.\ell$  to refer to the value of  $\ell$  in  $r$ . Similarly, if  $T$  is a record type and  $\ell$  is a label in  $T$ ,  $T.\ell$  refers to the type of  $\ell$  in  $T$ . Records (and record types) can be nested, so that the value of a label is itself a record (or record type). As can be seen in Figure 2, types can be constructed from predicates, e.g., “run” or “man”. Such types are called *ptypes* and correspond roughly to propositions in first order logic. Given a set of predicates and a set of possible arguments, the set of possible ptypes is **PType**, thus allowing for polymorphic predicates. The arity of a ptype  $P$  is a set of tuple of types  $\text{Arity}(P)$ . For example  $\text{Arity}(\text{run}) = \{\langle \text{Ind} \rangle\}$ .

A fundamental type-theoretical intuition is that something of a ptype  $T$  is whatever it is that counts as a proof of  $T$ . One way of putting this is that “propositions are types of proofs”. In Figure 2, we simply use  $\text{prf}(T)$  as a placeholder for proofs of  $T$ ; below, we will show how low-level perceptual input can be included in proofs.<sup>1</sup>

### 4 The left-or-right game

As an illustration, we follow Larsson (2015) in using a simple dialogue game called the left-or-right (LoR) game. In this game, one agent places objects on a square surface, and the other agent classifies these objects as being to the right or not. In first language acquisition, training of perceptual meanings typically takes place in situations where the referent is in the shared focus of attention and thus perceivable to the dialogue participants. We assume that our DPs (dialogue participants) are able to establish a shared focus of attention. A (simple) sensor collects some information (sensor input) from the environment and emits a real-valued vector. The sensor is assumed to be oriented towards the object in shared focus of attention.

### 5 Classifiers and TTR

Again following Larsson (2015), we formalise the notion of a simple perceptron classifier and provide its TTR type. The input to the classifier func-

<sup>1</sup>Note that TTR is not proof-theoretic like many other type theories. TTR proofs are more like *witnesses* in situation semantics (Barwise and Perry, 1983) or the *proof objects* in intuitionistic type theory (Martin-Löf and Sambin, 1984). For instance, there are no canonical proofs in TTR; there can be several non-equivalent proofs of the same ptype. This is related to the fact that types in TTR are intensional, i.e., there can be several different types with the same extension. Also, there is no notion of a proof method in TTR.

$$\left[ \begin{array}{l} \ell_1 = a_1 \\ \ell_2 = a_2 \\ \dots \\ \ell_n = a_n \\ \dots \end{array} \right] : \left[ \begin{array}{l} \ell_1 : T_1 \\ \ell_2 : T_2(\ell_1) \\ \dots \\ \ell_n : T_n(\ell_1, \ell_2, \dots, \ell_{n-1}) \end{array} \right]$$

Figure 1: Schema of record and record type

$$\left[ \begin{array}{l} \text{ref} = \text{obj}_{123} \\ c_{\text{man}} = \text{prf}(\text{man}(\text{obj}_{123})) \\ c_{\text{run}} = \text{prf}(\text{run}(\text{obj}_{123})) \end{array} \right] : \left[ \begin{array}{l} \text{ref} : \text{Ind} \\ c_{\text{man}} : \text{man}(\text{ref}) \\ c_{\text{run}} : \text{run}(\text{ref}) \end{array} \right]$$

Figure 2: Sample record and record type

tion  $\pi_{\text{right}}$  is (1) a parameter record specifying a weight vector  $w$  (a vector of real numbers) and a threshold  $t$  (a real number) and (2) a situation record specifying an object in the focus of attention,  $\text{foo}$ , and a sensor reading  $\text{sr}$  (a vector of real numbers). Whereas a (non probabilistic) classifier normally gives a Boolean output (corresponding to whether the neuron triggers or not), we want as output a ptype (or the negation thereof). The argument of the ptype predicate (right) is the object in the shared focus of attention, i.e. the value of the field  $\text{foo}$  in the situation record.

$$(1) \pi_{\text{right}} : \left[ \begin{array}{l} w : \mathbb{R}^+ \\ t : \mathbb{R} \end{array} \right] \rightarrow \left[ \begin{array}{l} \text{foo} : \text{Ind} \\ \text{sr} : \mathbb{R}^+ \end{array} \right] \rightarrow \text{Type}$$

such that if

$$\begin{aligned} &\bullet \text{par} : \left[ \begin{array}{l} w : \mathbb{R}^+ \\ t : \mathbb{R} \end{array} \right] \text{ and} \\ &\bullet r : \left[ \begin{array}{l} \text{foo} : \text{Ind} \\ \text{sr} : \mathbb{R}^+ \end{array} \right], \end{aligned}$$

then  $\pi_{\text{right}}(\text{par}, r) =$

$$\left\{ \begin{array}{ll} \text{right}(r.\text{foo}) & \text{if } r.\text{sr} \cdot \text{par}.w > \text{par}.t \\ \neg \text{right}(r.\text{foo}) & \text{otherwise} \end{array} \right.$$

Note that the function itself is defined outside TTR. This allows any classifier to be used with TTR, no matter how it is implemented. Classifiers can also be non-binary, as shown here for a fruit classifier FC:

$$(2) \pi_{\text{fruit}} : \mathbb{R}^+ \rightarrow \left[ \begin{array}{l} \text{foo} : \text{Ind} \\ \text{img} : \text{Image} \end{array} \right] \rightarrow \text{Type}$$

such that if

$$\bullet \text{par} : \mathbb{R}^+ \text{ and}$$

$$\bullet r : \left[ \begin{array}{l} \text{foo} : \text{Ind} \\ \text{img} : \text{Image} \end{array} \right],$$

then  $\pi_{\text{fruit}}(\text{par}, r) =$

$$\left\{ \begin{array}{ll} \text{apple}(r.\text{foo}) & \text{if FC}(r.\text{img}, \text{par}) = \text{Apple} \\ \text{orange}(r.\text{foo}) & \text{if FC}(r.\text{img}, \text{par}) = \text{Orange} \\ \text{pear}(r.\text{foo}) & \text{if FC}(r.\text{img}, \text{par}) = \text{Pear} \\ \dots & \\ \neg \text{fruit}(r.\text{foo}) & \text{otherwise} \end{array} \right.$$

## 6 Putting classification at the core

In this section, we present a version of TTR which explicitly puts classifiers at the core of what it is to understand natural language in relation to a perceived situation. This version replaces that of (Larsson, 2015) types and gives a clearer and more perspicuous account of how judgement and classification are related.

### 6.1 Meanings for predicates

We start by accounting for predicate meanings in TTR. Several types of expressions in natural language (nouns, verbs, adjectives) can be modelled semantically using predicates. We will represent the (perceptual) meaning of predicates as records containing four fields:

- Classifier parameters (params): a (possibly empty) record containing classifier parameters (e.g. weight vectors)
- Background meaning (bg): a record type representing assumptions about the context of utterance (presuppositions)
- Interpretation function (interp), taking a situation of type  $\text{bg}$  and providing a ptype encoding a contextual interpretation of an utterance in the context of that situation

- Classification function (clfr) that can be used to make a judgement as to whether an (interpreted) utterance correctly describes a situation

Accordingly, we define the type *Mng* of a meaning entry as follows:

$$(3) \text{ Mng} = \left[ \begin{array}{l} \text{params} : \text{Rec} \\ \text{bg} : \text{RecType} \\ \text{intrp} : \text{bg} \rightarrow \text{Type} \\ \text{clfr} : \text{bg} \rightarrow \text{Type} \end{array} \right]$$

Predicate meanings are defined for a predicate with a certain arity. It is convenient to have a looking function outputting the meaning of the predicate used in a given ptype. We define such a function *Pred* as follows (where  $P(a_1, \dots, a_n)$  is a ptype,  $P(a_1, \dots, a_n) \in \mathbf{PTYPE}$ ):

$$(4) \text{ Pred}(P(a_1, \dots, a_n)) = P_{\langle T_1, \dots, T_n \rangle}$$

where

- $\langle T_1, \dots, T_n \rangle \in \text{Arity}(P)$
- $a_1 : T_1, \dots, a_n : T_n$

For example, we get:

$$(5) \text{ Pred}(\text{right}(\text{obj}_{45})) = \text{right}_{\langle \text{Ind} \rangle}$$

Next, we define a function *PredMng* for looking up the meaning of a predicate, whose domain is  $\{P_A \mid P \in \mathbf{Pred}, A \in \text{Arity}(P)\}$  and whose range is in  $\{r \mid r : \text{Mng}\}$ . For example,

$$(6) \text{ PredMng}(\text{right}_{\langle \text{Ind} \rangle}) = \left[ \begin{array}{l} \text{params} = \left[ \begin{array}{l} \text{w} = [0.800 \quad 0.010] \\ \text{t} = 0.090 \end{array} \right] \\ \text{bg} = \left[ \begin{array}{l} \text{sr}_{\text{pos}} : \mathbb{R}^+ \\ \text{foo} : \text{Ind} \end{array} \right] \\ \text{intrp} = \lambda r : \text{bg} \cdot \text{right}(r.\text{foo}) \\ \text{clfr} = \lambda r : \text{bg} \cdot \pi_{\text{right}}(\text{params}, r) \end{array} \right]$$

We also define the interpretation of “right”:

$$(7) \llbracket \text{right} \rrbracket = \text{PredMng}(\text{right}_{\langle \text{Ind} \rangle}).\text{intrp}$$

Finally, we define

$$(8) \text{ Clfr}(T) = \text{PredMng}(\text{Pred}(T)).\text{clfr}$$

For example,

$$(9) \text{ Clfr}(\text{right}(\text{obj}_{45})) = \lambda r : \left[ \begin{array}{l} \text{sr}_{\text{pos}} : \mathbb{R}^+ \\ \text{foo} : \text{Ind} \end{array} \right] \cdot \pi_{\text{right}} \left( \left[ \begin{array}{l} \text{w} = [0.800 \quad 0.010] \\ \text{t} = 0.090 \end{array} \right], r \right)$$

## 6.2 Classification and witness conditions

We now get to the crux of how to put classifiers at the heart of our semantics. According to (Cooper, in progress), for  $T \in \mathbf{PTYPE}$ ,

$$(10) s : T \text{ iff } s \in F(T)$$

where  $F(T)$  is the *witness cache*, for type  $T$  – a set of situations (in the case of ptypes) previously judged to be of type  $T$ . The witness cache for a type and an agent can represent the history of judgements made by that agent with respect to the type.

We modify this definition to include witness conditions along the lines of PyTTR (Cooper, 2019) defined with respect to the classifier associated with the predicate of the ptype:

$$(11) s : T \text{ iff } \text{Clfr}(T)(s) = T \text{ or } s \in F(T)$$

This definition puts classifiers at the core of TTR. New judgements are made using the *Clfr* function. Previous judgements can be stored in the witness cache for  $T$ .

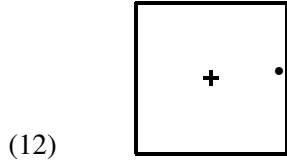
One issue that arises is in what to do first: apply the classifier, or check the witness cache? We do not take a stand on this issue here, but we note that checking the witness cache first makes sense provided it can be assumed to be up to date. Given that classifiers can be continuously trained on new instances, previous judgements may no longer be valid (in the sense that if they were made using the retrained classifier, the results would be different). Guaranteeing the validity of the witness cache would require that any changes in the classifier(s) related to a type  $T$  result in purging or re-evaluating the history of potentially affected judgements stored in the witness cache.

## 7 Putting the model to work

In this section, we show an illustrative example of how the framework above might be put to work in the context of the LoR game, when contextually interpreting utterances and when deciding whether they describe the situation correctly.

### 7.1 Interpretation

Assume that an agent  $A$  places an object on the surface and says “That one is to the right”, or just “Right”.



(12)

Agent  $B$  watches and gets a position sensor reading  $[0.900 \ 0.100]$  which is part of  $B$ 's take on the current situation ( $s_1$ ):

$$(13) \ s_1 = \begin{bmatrix} \text{sr}_{\text{pos}} & = & [0.900 \ 0.100] \\ \text{foo} & = & \text{obj}_{45} \end{bmatrix}$$

$B$  now interprets  $A$ 's utterance in the context the situation  $s_1$  by computing  $\llbracket \text{right} \rrbracket(s_1)$ , which gives the result  $\llbracket \text{right} \rrbracket(s_1) = \text{right}(\text{obj}_{45})$ . How does this happen? Recall that  $\llbracket \text{right} \rrbracket = \text{PredMng}(\text{right}_{\langle \text{Ind} \rangle}).\text{intrp}$ , which means that

$$(14) \ \llbracket \text{right} \rrbracket(s_1) = \\ (\text{PredMng}(\text{right}_{\langle \text{Ind} \rangle}).\text{intrp})(s_1) = \\ (\lambda r : \begin{bmatrix} \text{sr}_{\text{pos}} : \mathbb{R}^+ \\ \text{foo} : \text{Ind} \end{bmatrix} \cdot \text{right}(r.\text{foo})) ( \\ \begin{bmatrix} \text{sr}_{\text{pos}} = [0.900 \ 0.100] \\ \text{foo} = \text{obj}_{45} \end{bmatrix} ) = \\ \text{right}(\text{obj}_{45})$$

## 7.2 Classification

Next,  $B$  decides if  $A$ 's utterance correctly describes (her take on) the situation, i.e. if

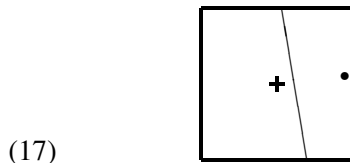
$$(15) \ s_1 : \llbracket \text{right} \rrbracket(s_1), \text{ i.e., if } s_1 : \text{right}(\text{obj}_{45})$$

For  $T = \text{right}(\text{obj}_{45})$ , we get

$$(16) \ s : \text{right}(\text{obj}_{45}) \text{ iff} \\ (\text{PredMng}(\text{right}_{\langle \text{Ind} \rangle}).\text{clfr})(s) = \text{right}(\text{obj}_{45}) \text{ or} \\ s \in F(\text{right}(\text{obj}_{45}))$$

In Figure 3, we show how this is checked for for  $s_1$ .

The result is that  $(\text{PredMng}(\text{right}_{\langle \text{Ind} \rangle}).\text{clfr})(s) = \text{right}(\text{obj}_{45})$ . Hence,  $s_1 : \text{right}(\text{obj}_{45})$  and (equivalently)  $s_1 : \llbracket \text{right} \rrbracket(s_1)$ . Consequently, this round of the LoR game plays out thus:



(17)

A: “right”  
B: “okay”

## 8 Vagueness and Probabilistic TTR

In Fernández and Larsson (2014), we formulate a Bayesian noisy threshold classifier for vague concepts such as “tall”. The classifier is trained on previous observations of tall entities, and is sensitive to the kind of entity (skyscraper, human, basketball player, ...). Instead of a binary judgement, the classifier returns an probability distribution over ptypes. This account connects to the probabilistic extension of TTR (Cooper et al., 2014, 2015).

Adapting from Fernández and Larsson (2014) to our current framework, the meaning of the vague predicate “tall” could be formalised thus:

$$(18) \ \text{PredMng}(\text{tall}_{\langle \text{Ind} \rangle}) \\ = \left[ \begin{array}{l} \text{bg} = \left[ \begin{array}{l} \text{c} : \text{Type} \\ \text{x} : \text{c} \\ \text{h} : \mathbb{R} \end{array} \right] \\ \text{params} = \left[ \begin{array}{l} \mu = \mu_{\text{tall}} \\ \sigma = \sigma_{\text{tall}} \end{array} \right] \\ \text{intrp} = \lambda r : \text{bg} \cdot \text{tall}(r.\text{x}) \\ \text{clfr} = \lambda r : \text{bg} \cdot \kappa_{\text{tall}}(\sigma(r.\text{bg}.\text{c}), \mu(r.\text{bg}.\text{c}), r.\text{h}) \end{array} \right] \\ \kappa_{\text{tall}} : (\mathbb{R}, \mathbb{R}, \text{bg}) \rightarrow [0, 1]$$

We are here employing a noisy *probabilistic threshold* (cf. Lassiter (2011)) – a normal random variable, represented by the parameters of its Gaussian distribution, the mean  $\mu$  and the standard deviation  $\sigma$  (the noise width). Note that the probabilistic threshold depend on the semantic class of the individual being classified:

$$(19) \ \mu_{\text{tall}} : \text{Type} \rightarrow \mathbb{R}$$

$$(20) \ \sigma_{\text{tall}} : \text{Type} \rightarrow \mathbb{R}$$

Interpretation works exactly as in the non-probabilistic case. Regarding classification, the probabilistic version of (11) above (ignoring the witness cache for the moment) is simply:

$$(21) \ p(s : T) = \text{Clfr}(T)(s)$$

Since the output of the clfr function is now a probability, so is the result of classification.

$$(22) \ p(s : \text{tall}(\text{sally})) \in [0, 1]$$

## 9 Conclusion

We presented a version of Type Theory with Records which places classifiers at the core of semantics. Using this framework, we present an account of the interpretation and classification of utterances referring to perceptually available information (such as a visual scene). The account improves on previous work by clarifying the role of

$$\begin{aligned}
& (\text{PredMng}(\text{right}_{(\text{Ind})}).\text{clfr})(s_1) \\
&= (\lambda r : \begin{bmatrix} \text{sr}_{\text{pos}} : \mathbb{R}^+ \\ \text{foo} : \text{Ind} \end{bmatrix} \cdot \pi_{\text{right}} \left( \begin{bmatrix} w = [0.800 & 0.010] \\ t = 0.090 \end{bmatrix}, r \right) \left( \begin{bmatrix} \text{sr}_{\text{pos}} = [0.900 & 0.100] \\ \text{foo} = \text{obj}_{45} \end{bmatrix} \right) \\
&= \pi_{\text{right}} \left( \begin{bmatrix} w = [0.800 & 0.010] \\ t = 0.090 \end{bmatrix}, \begin{bmatrix} \text{sr}_{\text{pos}} = [0.900 & 0.100] \\ \text{foo} = \text{obj}_{45} \end{bmatrix} \right) \\
&= \begin{cases} \text{right}(\text{obj}_{45}) & \text{if } [0.900 \quad 0.100] \cdot [0.800 \quad 0.010] > 0.090 \\ \neg \text{right}(\text{obj}_{45}) & \text{otherwise} \end{cases} \\
&= \text{right}(\text{obj}_{45})
\end{aligned}$$

Figure 3: Example classification derivation

classifiers in a hybrid semantics combining statistical/neural classifiers with logical/inferential aspects of meaning. The account covers both discrete and probabilistic classification, thereby enabling learning, vagueness and other non-discrete linguistic phenomena.

This account is intended as a starting point for a comprehensive account of semantics encompassing both referential and inferential meaning. Issues to explore include e.g. how referential meanings are coordinated between DPs, and how compositionality works for referential meaning (Larsson, 2017).

## Acknowledgements

This work was supported by grant 2014-39 from the Swedish Research Council (VR) for the establishment of the Centre for Linguistic Theory and Studies in Probability (CLASP) at the University of Gothenburg.

## References

- Jacob Andreas, Marcus Rohrbach, Trevor Darrell, and Dan Klein. 2016. Learning to compose neural networks for question answering. *arXiv preprint arXiv:1601.01705*.
- J. Barwise and J. Perry. 1983. *Situations and Attitudes*. The MIT Press.
- Robin Cooper. 2005. Austinian truth, attitudes and type theory. *Research on Language and Computation*, 3:333–362.
- Robin Cooper. 2012. Type theory and semantics in flux. In Ruth Kempson, Nicholas Asher, and Tim Fernando, editors, *Handbook of the Philosophy of Science*, volume 14: Philosophy of Linguistics. Elsevier BV. General editors: Dov M. Gabbay, Paul Thagard and John Woods.

Robin Cooper. 2019. Types as learnable cognitive resources in pytr. In Cleo Condoravdi and Tracy Holloway King, editors, *Tokens of Meaning: Papers in Honor of Lauri Karttunen*, pages 569–586. CSLI Publications.

Robin Cooper. in progress. *Type theory and language - From perception to linguistic communication*.

Robin Cooper, Simon Dobnik, Shalom Lappin, and Staffan Larsson. 2014. A probabilistic rich type theory for semantic interpretation. In *Proceedings of the EACL Workshop on Type Theory and Natural Language Semantics (TTNLS)*.

Robin Cooper, Simon Dobnik, Staffan Larsson, and Shalom Lappin. 2015. Probabilistic type theory and natural language semantics. *LiLT (Linguistic Issues in Language Technology)*, 10.

Simon Dobnik and Robin Cooper. 2013. **Spatial descriptions in type theory with records**. In *Proceedings of IWCS 2013 Workshop on Computational Models of Spatial Language Interpretation and Generation (CoSLI-3)*, pages 1–6, Potsdam, Germany. Association for Computational Linguistics.

Simon Dobnik, Staffan Larsson, and Robin Cooper. 2011. Toward perceptually grounded formal semantics. In *Proceedings of the Workshop on Integrating Language and Vision at NIPS 2011*, Sierra Nevada, Spain. Neural Information Processing Systems Foundation (NIPS).

Raquel Fernández and Staffan Larsson. 2014. Vagueness and learning: A type-theoretic approach. In *Proceedings of the 3rd Joint Conference on Lexical and Computational Semantics (\*SEM 2014)*, Dublin, Ireland. The \*SEM 2014 Organizing Committee.

Mehdi Ghanimifard and Simon Dobnik. 2017. Learning to compose spatial relations with grounded neural language models. In *IWCS 2017-12th International Conference on Computational Semantics-Long papers*.

- Stevan Harnad. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1990):335–346.
- Casey Kennington and David Schlangen. 2015. Simple learning and compositional application of perceptually grounded word meanings for incremental reference resolution. In *Proceedings of the Conference for the Association for Computational Linguistics (ACL)*, pages 292–301.
- Staffan Larsson. 2011. The ttr perceptron: Dynamic perceptual meanings and semantic coordination. In *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue (SemDial 2011)*, Los Angeles (USA). Institute for Creative Technologies.
- Staffan Larsson. 2015. Formal semantics for perceptual classification. *Journal of Logic and Computation*, 25(2):335–369. Published online 2013-12-18.
- Staffan Larsson. 2017. Compositionality for perceptual classification. In *IWCS 2017—12th International Conference on Computational Semantics—Short papers*.
- Staffan Larsson and Robin Cooper. 2009. Towards a formal view of corrective feedback. In *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition, EACL*, pages 1–9, Athens, Greece. Association for Computational Linguistics.
- Staffan Larsson and Jenny Myrendal. 2017. Dialogue acts and updates for semantic coordination. In *Proc. SEMDIAL 2017 (SaarDial) Workshop on the Semantics and Pragmatics of Dialogue*, pages 52–59.
- Dan Lassiter. 2011. [Vagueness as probabilistic linguistic knowledge](#). In R. Nowen, R. van Rooij, U. Sauerland, and H. C. Schmitz, editors, *Vagueness in Communication*. Springer.
- Diego Marconi. 1997. *Lexical competence*. MIT press.
- P. Martin-Löf and G. Sambin. 1984. *Intuitionistic type theory*. Studies in proof theory. Naples: Bibliopolis.
- Richard Montague. 1974. *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press, New Haven. Ed. and with an introduction by Richmond H. Thomason.
- David Schlangen, Sina Zarriess, and Casey Kennington. 2016. Resolving References to Objects in Photographs using the Words-As-Classifiers Model. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)*.
- Todd Shore and Gabriel Skantze. 2018. Using lexical alignment and referring ability to address data sparsity in situated dialog reference resolution. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2288–2297.
- Ronja Utescher. 2019. Visual ttr-modelling visual question answering in type theory with records. In *Proceedings of the 13th International Conference on Computational Semantics-Student Papers*, pages 9–14.