ERC Starting Grant 2016
Research proposal [Part B1]

A distributional MOdel of Reference to Entities AMORE

Cover Page:
- Name of the Principal Investigator (PI): Gemma Boleda
- Name of the PI's host institution for the project: Universitat Pompeu Fabra
- Proposal duration in months: 60

When I asked my seven-year-old daughter "Who is the boy in your class who was also new in school last year, like you?", she instantly replied "Daniel", using the descriptive content in my utterance to identify an entity in the real world and refer to it. The ability to use language to refer to reality is crucial for humans, and yet it is very difficult to model. AMORE breaks new ground in Computational Linguistics, Linguistics, and Artificial Intelligence by developing a model of linguistic reference to entities implemented as a computational system that can learn its own representations from data.

This interdisciplinary project builds on two complementary semantic traditions: 1) Formal semantics, a symbolic approach that can delimit and track linguistic referents, but does not adequately match them with the descriptive content of linguistic expressions; 2) Distributional semantics, which can handle descriptive content but does not associate it to individuated referents. AMORE synthesizes the two approaches into a unified, scalable model of reference that operates with individuated referents and links them to referential expressions characterized by rich descriptive content. The model is a distributed (neural network) version of a formal semantic framework that is furthermore able to integrate perceptual (visual) and linguistic information about entities. We test it extensively in referential tasks that require matching noun phrases ("the Medicine student", "the white cat") with entity representations extracted from text and images.

AMORE advances our scientific understanding of language and its computational modeling, and contributes to the far-reaching debate between symbolic and distributed approaches to cognition with an integrative proposal. I am in a privileged position to carry out this integration, since I have contributed top research in both distributional and formal semantics.

Section a: Extended Synopsis of the scientific proposal

Background and goals

Humans use language to communicate about the world, from immediate sensory information ("Caution, it's hot!") to very abstract information acquired through the years or even through generations ("The Universe is expanding"). Modeling the process by which we use linguistic expressions to refer to the outside world is fundamental for understanding language, a defining trait of the human species. The goal of AMORE is to advance the state of the art in Computational Linguistics, Linguistics, and Artificial Intelligence, by developing **a model of linguistic reference to entities implemented as a computational system** that can learn its own representations from data. I focus on concrete entities (physical objects humans perceive as a unit) because they constitute a well-delimited domain, representative of the larger reference problem (see *Feasibility* below).

Linguistic reference crucially involves both **fuzzy** and **discrete** aspects of meaning. A noun phrase such as "the big tree" gives us some *descriptive content* that allows us to identify a particular entity through some of its properties (Frege 1892). However, this descriptive content is notoriously fuzzy and vague (Fodor et al. 1980; Cruse 1986; Keefe 2000): The word "tree" applies to "many unlike individuals of diverse size and form" (Borges 1944), from near-bushes to sequoyas to even genealogical trees, with no definite criteria nor clear boundaries between what counts as a tree and what doesn't. Fuzziness persists when composing words into phrases and sentences: For instance, the meaning of "red" straddles towards e.g. pink or orange depending on the modified noun (compare "red car" and "red cheek"; Boleda et al. 2013). The fuzzy nature of meaning is actually a very useful trait of language, and the conceptual system it relies on, because it allows us to handle an infinitely varied, ever-changing reality reusing knowledge about previously encountered situations (Murphy 2002; van Deemter 2010).

When phrases with fuzzy descriptive content are used in a specific context, however, they are used to refer to specific entities in the real world. Humans treat the *referents* picked out by phrases as essentially discrete, and language offers us tools to deal with that, too. For instance, a speaker uttering example (1) uses the noun phrase "a box" to introduce a referent in the current discourse, and the discrete anaphoric pronoun "it" to refer back to precisely that referent and add more information about it (Kamp and Reyle 1993).

(1) The man lifted a box. It was heavy.

Thus, linguistic referents are both discrete (they are *individuated* through linguistic mechanisms such as the use of noun phrases and pronouns) and fuzzy (they are linked to rich descriptive content). Some of the most successful previous work in theoretical and computational semantics, however, is markedly biased towards one aspect, at the expense of the other. *Formal semantics* (Montague 1970 and subsequent work) employs logic and other symbolic mathematical tools to provide discrete semantic representations of linguistic expressions. This approach has advanced our understanding of the linguistic mechanisms that individuate referents. For instance, it models the fact that when I say "the big tree" I pick out a unique object that is a big tree, that indefinite noun phrases ("a box") are used to introduce new referents in the discourse whereas definite ones ("the man") point to already accessible referents, and that pronouns are pointers to referents (Kamp and Reyle 1993). However, formal semantics says little about the characteristic properties of a big tree or a box, and consequently about how we are able to pick out the right entity in the first place. *Distributional semantics* (see Turney and Pantel 2010 for an overview) models meaning in terms of context of use, because semantically related expressions are used in similar contexts (Harris 1968). This framework provides numerical, continuous distributed representations for linguistic expressions that are useful to model descriptive content: Distributional semantics models graded semantic phenomena such as word and phrase similarity ("box" – "package", "important route" – "major road"; Landauer and Dumais 1997; Baroni and Zamparelli 2010) and meaning modulation (the difference between red in "red car" vs. "red cheek"; Boleda et al. 2013). However, current distributional models do not have a notion of an individuated referent to attach the descriptive content to: Their semantic representations do not have the right structure to distinguish between different referents. As a result, they might for instance tell us that "box" and "package", being conceptually similar, can in principle denote the same referent, but they are useless at telling us if, in a specific discourse, the two expressions are indeed pointing to the same referent or not.

The shortcomings of each approach severely limit their use in Computational Linguistics tasks that require natural language interpretation. For instance, Bos and Markert (2005) applied a formal semantic system to Recognizing Textual Entailment, which seeks to model natural language inference (e.g., "Crude oil prices soared to record levels" entails "Crude oil prices rose", but "A white man spoke" does not entail "A black man spoke"). The system was reasonably precise: When it predicted entailment, it was right 77% of the time. However, it was able to predict only 6% of the entailments; because of its poor treatment of descriptive content, it had very low coverage. This system was e.g. unable to relate "[a] sport utility vehicle drifted onto the shoulder of a highway and struck a parked truck" to "car accident". Systems that solely use distributional semantics have a better coverage but a lower precision (Beltagy et al. 2013): Entailment decisions require predicting whether an expression applies to the same referent or not, and these systems make trivial mistakes, such as deciding that "white man" and "black man", being conceptually similar, should co-refer.

I believe that the lack of tools to adequately handle descriptive content is an insurmountable roadblock for symbolic approaches such as formal semantics. My strategy therefore is to take inspiration from some important insights in the formal tradition, but focus on pushing the limits of distributional semantics, a fast-moving and exciting area to which I am strongly contributing. AMORE will **endow distributional semantics with referential capabilities**. Thanks to my original adaptation to language of very recent developments in Machine Learning, the model will be able to automatically induce and operate with **individuated referents** which, however, have **a continuous distributed internal representation**, accounting for the conceptual richness of the process of referring. The use of these advances further allows the model to integrate two major sources of information about referred entities (Kamp 2015): **Perceptual** information from the environment, and **previous knowledge** gained through language.

The goals of AMORE are to:

- Develop a model of reference to entities that links descriptive content and individuated referents.
- Implement it as a computational (specifically, neural network) system that is able to learn representations from data.
- Test it extensively in experiments involving reference to entities in discourse and in the perceptual (visual) environment.
- Explore the consequences of moving to distributed entity representations for Computational Linguistics, Linguistics more generally, and Artificial Intelligence.

Methodological foundations

Distributional semantics represents linguistic expressions with *vectors* (essentially lists of numbers; they can also be more complex algebraic objects such as matrices and tensors). For instance, to represent word meaning, each word is assigned a vector with a large number of dimensions or features. The semantic information is *distributed* across all the dimensions of the vector, and is expressed in the form of *continuous* values. This allows for rich and nuanced information to be encoded for each word (Landauer and Dumais 1997; Baroni and Zamparelli 2010). Semantically related words have similar vectors, because their values are *learned* from natural language data. Specifically, vector values are abstractions on the *contexts* in which words are used, since related words, such as "box" and "package", are used in similar contexts ("open the _", "a light _"; Harris 1968). In recent years, distributional models have been extended to handle the semantic *composition* of words into phrases and sentences (Mitchell and Lapata 2010, Socher et al. 2013, Baroni et al. 2014a). Although these models still do not account for the full range of composition phenomena that have been examined in formal semantics, they do encode relevant semantic information, as shown by their success in demanding semantic tasks such as predicting sentence similarity (Marelli et al. 2014). Another recent research direction is the addition of perceptual information to semantic representations, e.g. using information extracted from images to improve the representation of color terms (Bruni et al. 2012).

Neural networks, a family of Machine Learning algorithms developed in the 1960s that are receiving renewed interest due to significant breakthroughs in many Artificial Intelligence areas (LeCun et al. 2015), can obtain better representations for linguistic expressions than traditional distributional semantic methods, by learning the abstraction operations to perform on the linguistic contexts using context prediction tasks (Mikolov et al. 2013, Baroni et al. 2014b). Moreover, such models naturally extend distributional learning to handle complex linguistic expressions, such as sentences and discourse chunks, and provide a principled supervised framework to adapt it to a variety of tasks (Li et al. 2015). AMORE exploits very recent progress in the field (e.g., Joulin and Mikolov 2015, Sukhbaatar et al. 2015, Bahdanau et al. 2015) to endow its network with a dynamic trainable memory able to handle individuated referents.

The model

The AMORE model is essentially a distributional version of Kamp (2015), a formal semantic model (based on Discourse Representation Theory; Kamp and Reyle 1993) that is one of the most comprehensive theoretical models of the interpretation of noun phrases to date. Kamp's model has four main components, accounting for: (1) generic information about the world, such as the fact that books have covers (K_{gen}); (2) information about the immediate extralinguistic environment (K_{env}); (3) the current linguistic discourse (K_{dis}); (4) the entities we talk about (K_{enc}, for "encyclopedic"). Interestingly, Kamp explicitly declares K_{gen} "off-limits" for his model (Kamp 2015, p. 54), consistent with the fact that it concerns descriptive content.

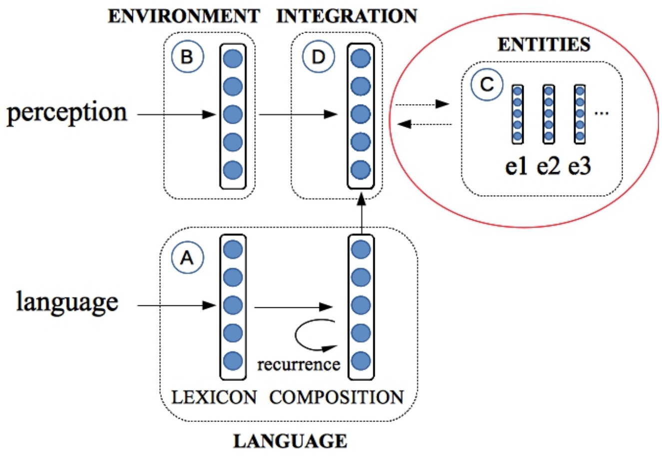

Figure 1. The AMORE model.

Figure 1 depicts the model to be developed in the project. It has four main components, labeled with letters in the figure. The first component (A) processes the *linguistic* input, and it includes a distributional word lexicon and a compositional representation of discourse meaning incrementally computed from the word representations; component (B) represents the perceptual *environment*; and component (C) handles the *entities* in the discourse and the environment. The fourth component (D) integrates visual and linguistic information and transfers information to and from the entity library.

The K_{gen} component of Kamp's model is operationalized in AMORE as the corpus-induced distributional lexicon, because distributional representations have been shown to account for generic information (Baroni and Lenci 2010, a.o.). K_{env} and K_{enc} correspond to our environment component and entity library, respectively. Both in Kamp's and in our model, the entity library contains representations for previously known entities (if I say to my mother "Aunt Lina came today", she can access her long-term entity representation for her sister) as well as those newly introduced in a discourse ("I just saw *a bird*"). This project focuses mainly on newly introduced entities (Work Packages 3.1, 4.1 and 4.2), but one of the experiments is targeted at previously known entities (WP 3.2). The information in K_{dis} is distributed in AMORE between the composition module, the integration module, and the entity library.

The model is implemented as a Recurrent Neural Network (RNN; Elman 1990), extended with a dynamic memory for the entity library (see below), an additional hidden layer for the integration component, and a pre-trained image-processing neural network for the environment component (Krizhevsky et al. 2012). We choose RNNs because they have some semantic composition capabilities and, especially when augmented with the gating mechanisms proposed by Hochreiter and Schmidhuber (1997), partially account for syntax (Li et al. 2015). Note that the compositional operations putting semantic representations together, as well as the relevant syntactic properties, are also learned by the model, rather than pre-specified.

A RNN induces a single continuous representation of the semantic information active at each time step in the unfolding discourse, with nothing akin to the distinct variables used in formal semantics to represent entities. The crucial technical innovation of AMORE (highlighted with a red ellipse in Figure 1) will be to extend RNNs with a dynamic memory that the network can learn to manipulate, storing distinct representations in it, and retrieving them when needed. This allows us to emulate **discrete operations on entity representations in a continuous setup**. The specific mechanisms build on recent work that uses continuous approximations to discrete operations such as storage and retrieval, in order to integrate them in architectures that, being fully differentiable, can learn from data through effective methods (e.g., Joulin and Mikolov 2015, Sukhbaatar et al. 2015, Bahdanau et al. 2015). As an example of what we expect the model to learn, consider input "a bird": The network should associate the use of the indefinite article "a" (which typically introduces new discourse referents) with the operation that inserts a new entity in the library, and the noun "bird" with the update operation that provides properties associated to that referent. Technical details are provided in B2.

The architecture just specified needs appropriate data to learn from. Note in particular that the entity library will not be manually encoded, but it will have to be constructed by the model motivated by the tasks it is faced with. The model is trained on tasks that encourage it to create and access entity representations (see WPs 3 and 4 below). For instance, in WP3.1 the task is to predict the name of the character in a story that corresponds to a given noun phrase: Given a story that starts "Bob is a Law student, but Ann studies Medicine", the system should output "Ann" in response to "The Medicine student". The specific implementation of the output-producing mechanism depends on the task (see WP2 below).

The AMORE model is thus **able to keep entities distinct** (they correspond to different vectors in the entity library) **while at the same time providing rich internal distributed representations for them**, since the descriptive content and perceptual features associated with the entity are stored in its distributed representation. We can thus capitalize on the power of distributional semantics to handle different linguistic expressions that are semantically related ("box" - "package") and integrate linguistic and perceptual information (linking noun phrases like "the white cat" to entities depicted in images); at the same time, we also emulate symbolic variables by letting the network interact with the entity library. This is a highly novel approach to entity representations, and the first linguistically motivated use of neural networks with dynamic memory structures.

Methodology

Work Package 1: Data and infrastructure. WP1.1 (Computational infrastructure) sets the necessary computational infrastructure, WP2.2 (Data/code release) makes data and models publicly available.

Work Package 2: Model development. WP2.1 (Model definition and implementation) develops the formal definition of the model and its implementation as a neural network with the architecture specified above. This includes also smart component initialization, e.g. using pre-trained representations for the distributional lexicon (Mikolov et al. 2013). WP2.2 (Task adaptation) defines how the network is adapted to the input and target output required by each task. For instance, for the "Ann" task in WP 3.1, the network first processes the whole story one word at a time. At query time, the network processes the noun phrase of interest (e.g., "The Medicine student") also one word at a time. The most related entity in the library is then "softly" retrieved by relying on the similarity of the phrase representation to all entity vectors in the library. A matching operation uses that entity to return the output name (in the example, "Ann") from the distributional lexicon. The architecture for the experiments with images (WP 4) is analogous, except that the input includes image representations presented alone or in parallel with the last word in the corresponding verbal descriptions. The task to identify characters in scripts (WP 3.2) will be operationalized as a standard language modeling ("predict the next word") task, as detailed in B2.

Work Package 3: Linguistic knowledge-based reference. This WP models reference to entities based on knowledge about them coming language (discourse). In WP3.1 (Reference to newly-introduced entities) we use the "Ann" task explained above: Given a story with several formerly unknown entities (characters, locations, etc.), decide which entity a given noun phrase refers to. We will collect a large data set of stories and associated referential expressions via crowdsourcing (a method that facilitates large scale data collection via paid tasks done through an online interface), and we consider the resulting publicly available corpus a major deliverable of the project. WP3.2 (Long-term modeling of entities) tests the ability of the model to handle previously known entities (recall the "Aunt Lina" example above), through a task that thoroughly probes the entity library across time. The task is to associate each utterance in the second part of a movie script with the character that produced it, based on the entity library built in the first part. For this task we can produce a large amount of training data automatically, by collecting and processing freely available movie scripts.

Work Package 4: Perception-based and integrated reference. This WP models reference using perceptual (visual) information, on its own or integrated with knowledge acquired through language. We ask the system to identify an entity depicted in an image through a noun phrase with descriptive content (e.g., using "the white cat" to identify a white cat in a picture). Because this is an ambitious task that requires the integration of all the components in the model, the experiments are progressively more difficult. Experiment 1 of WP4.1 (Controlled domain) focuses on visual properties. We will present the system with a sequence of images (for instance, a plant, a white cat, a black dog, a black cat, a white dog) followed by one noun phrase. The task of the system is to decide whether it has seen a unique entity corresponding to that noun phrase. In the example, the system should reply "yes" to "white cat", "no" to "white animal" (because it has seen two), and "no" to "brown dog". Note that since multiple images (objects) are presented, and the model is tested after their sequential presentation, this and the following tasks are crucially probing the model's ability to store representations of different entities in its library, keeping them distinct through time. We will use a pre-existing data set (Russakowski and Li 2012) with images of single objects annotated with visual attributes, using the information in the dataset to associate each image with descriptive noun phrases. Experiment 2 in WP4.1 adds further knowledge about the entities coming from language: We will associate each image in the input with linguistic expressions (e.g., adding "is sociable" to the white cat image) and evaluate the model on noun phrases using only visual information ("white cat"), only language-based knowledge ("is sociable"), and an integration of the two ("sociable white cat", "sociable animal"). The linguistic expressions will be harvested from available textual corpora. WP 4.2 (Open domain) tackles the same tasks as WP4.1 in a much more challenging setting: Naturalistic images with human-produced referring expressions (using the ReferItGame dataset, Kazemzadeh et al. 2014).

Work Package 5: Coordination and dissemination. Activity 1 (project management) monitors scientific and organizational progress, including consulting with an external Advisory Board consisting of renowned experts in distributional and formal semantics (Katrin Erk, Hans Kamp, Hinrich Schütze). Activity 2 (dissemination) foresees publications, talks, a dedicated project webpage, and society outreach activities.

Team

The AMORE team consists of myself (75% dedication), senior member Louise McNally (10% dedication), three post-doctoral researchers, and two PhD students. I will lead the project. L. McNally, an expert in formal semantics and long-term collaborator, will participate in the formal specification of the model (WP 2) and in assessing the implications of the project for theoretical linguistics. Post-doc 1 (with expertise in neural networks) will develop and implement the model (WP 2). Post-doc 2 (with expertise in distributional semantics and/or discourse) and PhD student 1 will carry out the experiments on linguistic knowledge-based reference (WP 3). Post-doc 3 (with expertise in integrating language and vision) and PhD student 2 will carry out the experiments on perception-based and integrated reference (WP 4).

High-risk high-gain nature of the project and feasibility

AMORE is a radically new approach to reference with high-gain potential, because it provides a unified, scalable way to deal with the fuzzy and discrete aspects of reference, bringing about a revolutionary step forward in theoretical and computational semantics. It is also high-risk, specifically with respect to the following points: (1) The model is completely data-induced, with little control (beyond specifying its architecture) over what each component does. This is what makes it powerful (it is able to learn), but it also makes it risky. In particular, there is no guarantee that the entity library will effectively store entity representations, as opposed to other types of information. (2) While distributional representations have been shown to be very effective to represent words and more complex linguistic expressions, there is no significant previous experience in using them for discourse entities. Given its inherent high-risk nature, I have carefully designed the project so as to maximize **feasibility**. First, given that the reference problem is huge, I have selected a well-defined domain with a solid theoretical framework to rely on: Reference to concrete entities (excluding e.g. abstract entities and events), focusing on single-entity denoting noun phrases (excluding e.g. plurals and group nouns). Second, I have designed tasks that will encourage the model to build entity representations and to match them with the representations obtained from the other components. Third, in the model I am using components that have previously been shown to work for tasks related to AMORE, in my research as well as that of others. Fourth, the experiments include simpler and more complex challenges, and will be informative even in case of partial success. Finally, even if the AMORE model itself fails, because the project poses challenges to the community that need solving and operationalizes them in a way that is at the limits of the state of the art, its results (data sets, modeling results) will be very valuable for the community. For instance, if the entity library stores other information, it could be revealing to analyze what it stores. My previous research experience also supports the feasibility of the project. I am in a privileged position to carry out this project, since I have contributed top research in both distributional semantics (Boleda et al. 2004, Mayol et al. 2005, Boleda et al. 2007, Boleda et al. 2012b,c,d, Bruni et al. 2012, Boleda et al. 2013, Roller et al. 2014, Boleda and Erk 2015, Gupta et al. 2015) and formal semantics (McNally and Boleda 2004, Boleda et al. 2012a, Arsenijević et al. 2014, McNally and Boleda 2015). I have carried out extensive research on specific topics that are relevant for AMORE, such as adjectival and nominal semantics and, recently (as the PI of a Marie Curie project), on comparing symbolic and distributed semantic representations, integrating visual and linguistic information, and extracting referential information from distributional vectors. I have also shown my leadership capabilities, for instance encouraging the cross-fertilization between formal and distributional semantics as a guest editor for a special issue on the topic in the top journal in Computational Linguistics and leading the development of computational linguistic resources (see CV).

Expected impact

AMORE is a highly interdisciplinary project pushing Linguistics, Computational Linguistics, and Artificial Intelligence forward. The proposed model handles reference, as does formal semantics, but in a framework that can adequately deal with descriptive content, integrates perceptual and linguistic information, and can learn from data and so has a broad coverage. These advances will contribute to our scientific understanding of language, a defining trait of the human species. The project also contributes to the decades-long debate in Artificial Intelligence and Cognitive Science over *symbolic* vs. *distributed* approaches to cognition (Fodor and Pylyshyn 1988, Churchland 1998, Fodor and Lepore 1999, LeCun et al. 2015, among many others) with a proposal that synthesizes strong aspects of both approaches. The proposal will impact Artificial Intelligence as the first linguistically interesting application of algorithms that have until now been used either for toy tasks (Joulin and Mikolov 2015) or without clear linguistic motivation (Weston et al. 2014). From a more applied perspective, although the project itself focuses on fundamental research, it paves the way for technologies that will enable machines to talk to us in situated applications.

REFERENCES

Antol, S., Agrawal, A., Lu, J., Mitchell, M., Batra, D., Zitnick, C. L., & Parikh, D. (2015). VQA: Visual Question Answering. *arXiv preprint arXiv:1505.00468*.

Arsenijevic, B., Boleda Torrent, G., Gehrke, B., & McNally, L. (2010). Ethnic adjectives are proper adjectives. In Baglini, R., Grinsell, T., Keane, J.A., Singerman, R. & Thomas, J. (Eds.), *CLS 46-I The Main Session: Proceedings of 46th Annual Meeting of the Chicago Linguistic Society*, 17–30.

Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. In *Proceedings of ICLR*, 1–15.

Baroni, M., & Lenci, A. (2010). Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics*, *36*(4), 673–721.

Baroni, M., & Zamparelli, R. (2010). Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of EMNLP*, 1183–1193.

Baroni, M., Bernardi, R., & Zamparelli, R. (2014a). Frege in space: A program of compositional distributional semantics. *Linguistic Issues in Language Technology*, *9*.

Baroni, M., Dinu, G., & Kruszewski, G. (2014b). Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. In *Proceedings of ACL*, 238–247.

Boleda, G., im Walde, S. S., & Badia, T. (2007). Modelling Polysemy in Adjective Classes by Multi-Label Classification. In *Proceedings of EMNLP-CoNLL*, 171–180.

Boleda, G., Evert, S., Gehrke, B., & McNally, L. (2012a). Adjectives as saturators vs. modifiers: Statistical evidence. In Aloni, M., Kimmelman, V., Roelofsen, F., Sassoon, G. W., Schulz, K. & Westera M. (Eds.) *Logic, Language and Meaning - 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, 112–121. Springer.

Boleda, G., Padó, S., & Utt, J. (2012b). Regular polysemy: A distributional model. In *Proceedings of *SEM*, 151–160.

Boleda, G., Vecchi, E. M., Cornudella, M., & McNally, L. (2012c). First-order vs. higher-order modification in distributional semantics. In *Proceedings of the EMNLP-CONLL*, 1223–1233.

Boleda, G., im Walde, S. S., & Badia, T. (2012d). Modeling regular polysemy: A study on the semantic classification of Catalan adjectives. *Computational Linguistics*, *38*(3), 575–616.

Boleda, G., Baroni, M., The Pham, N. & McNally, L. (2013). Intensionality was only alleged: On adjective-noun composition in distributional semantics. In *Proceedings of IWCS*, 35–46.

Boleda, G., & Erk, K. (2015). Distributional Semantic Features as Semantic Primitives—Or Not. In *Knowledge Representation and Reasoning: Integrating Symbolic and Neural Approaches. Papers from the AAAI Spring Symposium*.

Beltagy, I., Chau, C., Boleda, G., Garrette, D., Erk, K., & Mooney, R. (2013). Montague meets Markov: Deep semantics with probabilistic logical form. *Proceedings of* SEM*, 11–21.

Borges, J. L. (1944). Funes El Memorioso. In *Ficciones (1935-1944)*. Ediciones SUR. Citing from the translation by J. E. Irby, 2000, Funes the Memorious.

Bos, J., & Markert, K. (2005). Recognising textual entailment with logical inference. In *Proceedings of HLT-EMNLP*, 628–635.

Bruni, E., Boleda, G., Baroni, M., & Tran, N. K. (2012). Distributional Semantics in Technicolor. In *Proceedings of ACL*, 136–145.

Churchland, P. M. (1998). Conceptual similarity across sensory and neural diversity: The Fodor/Lepore challenge answered. *The Journal of Philosophy*, 5–32.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge University Press.

van Deemter, K. (2010). *Not exactly: In praise of vagueness*. Oxford University Press.

Elman, J. L. (1990). Finding structure in time. *Cognitive science*, *14*(2), 179–211.

Fodor, J. A., Garrett, M. F., Walker, E. C., & Parkes, C. H. (1980). Against definitions. *Cognition*, *8*(3), 263–367.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and Cognitive Architecture: a Critical Analysis. *Cognition*, 28: 3–71.

Fodor, J., & Lepore, E. (1999). All at sea in semantic space: Churchland on meaning similarity. *the Journal of Philosophy*, 381–403.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, *100*, 25–50.

Gupta, A., Boleda, G., Baroni, M., & Padó, S. (2015). Distributional vectors encode referential attributes. In *Proceedings of EMNLP*, 12–21.

Harris, Z. (1968). *Mathematical Structures of Language*. Wiley.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*8), 1–32.

Joulin, A., & Mikolov, T. (2015). Inferring Algorithmic Patterns with Stack-Augmented Recurrent Nets. *arXiv preprint arXiv:1503.01007*.

Kamp, H. (2015). Entity Representations and Articulated Contexts An Exploration of the Semantics and Pragmatics of Definite Noun Phrases. Ms., University of Stuttgart.

Kamp, H., & Reyle, U. (1993). *From discourse to logic*. Kluwer.

Kazemzadeh, S., Ordonez, V., Matten, M., & Berg, T. L. (2014). ReferItGame: Referring to Objects in Photographs of Natural Scenes. In *Proceedings of EMNLP*, 787–798.

Keefe, R. (2000). *Theories of vagueness*. Cambridge University Press.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

LeCun, Y., Bengio, Y. & Hinton, G. E. (2015) Deep Learning. *Nature*, *521*, 436–444.

Li, J., Jurafsky, D., & Hovy, E. (2015). When Are Tree Structures Necessary for Deep Learning of Representations? In *Proceedings of EMNLP*, 2304–2314.

Marelli, M., Bentivogli, L., Baroni, M., Bernardi, R., Menini, S., & Zamparelli, R. (2014). SemEval-2014 Task 1: Evaluation of compositional distributional semantic models on full sentences through semantic relatedness and textual entailment. In *Proceedings of SemEval*, 1–8.

Mayol, L., G. Boleda & Badia, T. (2005). Automatic acquisition of syntactic verb classes with basic resources. *Language Resources and Evaluation*, 39(4), 295–312.

McNally, L. & Boleda, G. (2004). Relational adjectives as properties of kinds. In Olivier Bonami and Patricia Cabredo Hofherr (eds.) *Empirical Issues in Syntax and Semantics 5*, 179–196.

McNally, L. & Boleda, G. (2015). Conceptual vs. Referential Affordance in Concept Composition. To appear in Winter, Y. & Hampton, Y. (eds.) *Concept Composition and Experimental Semantics/Pragmatics*. Springer.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, 3111–3119.

Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science*, *34*(8), 1388–429.

Montague, R. (1970). English as a formal language. In Bruno Visentini (Ed.) *Linguaggi nella società e nella tecnica*. Mailand 1970, 189–223. Reprinted in (Thomason 1974), 188–221..

Murphy, G. L. (2002). *The big book of concepts*. MIT press.

Roller, S., Erk, K., & Boleda, G. (2014). Inclusive yet selective: Supervised distributional hypernymy detection. In *Proceedings of COLING*, 1025–1036.

Russakovsky, O., & Fei-Fei, L. (2012). Attribute learning in large-scale datasets. In *Trends and Topics in Computer Vision*, 1–14. Springer.

Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of EMNLP*, 1631–1642.

Sukhbaatar, S., Szlam, A., Weston, J., & Fergus, R. (2015). End-To-End Memory Networks, 1–11. *arXiv preprint arXiv*:1503.08895

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, *37*(1), 141–188.

Thomason, R. H. (Ed.) (1974). *Formal Philosophy, Selected Papers of Richard Montague*. Yale University Press.

Weston, J., Chopra, S., & Bordes, A. (2014). Memory networks. *arXiv preprint arXiv:1410.3916*.

Witten, I. H., & Frank, E. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.

Section b: Curriculum Vitae

Periodically updated information at: http://gboleda.utcompling.com.

EDUCATION

2007	PhD Universitat Pompeu Fabra, Universitat Pompeu Fabra, Spain.
2003	Master Cognitive Science and Language, Universitat Pompeu Fabra, Spain (with Honors).
2000	B. A. in Spanish Philology, Universitat Autònoma de Barcelona, Spain (with Honors).

EMPLOYMENT

2015 – pres.	Post-doctoral researcher*, University of Trento, Italy.
2014 – 15	Post-doctoral researcher* and lecturer, Universitat Pompeu Fabra, Spain.
2012 – 14	Post-doctoral researcher* and lecturer, The University of Texas at Austin, USA.
2011 – 12	Researcher, Universitat Pompeu Fabra, Spain.
2010	Visiting researcher*, University of Stuttgart, Germany.
2008 – 11	Post-doctoral researcher*, Universitat Politècnica de Catalunya, Spain.
2005 – 07	Researcher, Barcelona Media Centre d'Innovació, Spain.
2001 – 07	Doctoral researcher*, Universitat Pompeu Fabra, Spain.
2004, 2003	Visiting researcher*, Saarland University, Germany.
2000	Student assistant, Universitat Pompeu Fabra, Spain.
2000	Student assistant, Artificial Intelligence Research Institute (IIIA, CSIC), Spain.
1997 – 98	Student assistant, University of Cologne, Germany.

**Positions funded through competitive fellowships; see Section c.*

PARTICIPATION IN FUNDED PROJECTS

As PI

***Note**: because of the short duration of post-doctoral contracts, I have not been able to apply for regular projects as a PI (see employment).*

2015 – 2017	EU, Marie Skłodowska-Curie Program, H2020-MSCA-IF-2014 655577 ("LOVe: Linking Objects to Vectors in distributional semantics: A framework to anchor corpus-based meaning representations to the external world"), €180,277.20

As collaborating researcher / member

EU	One FP6 project (€1,000,000), one Network of Excellence
USA	One DARPA grant ($1,302,318)
Spain	Nine projects, one Network of Excellence
Catalonia	Two funded research groups

SUPERVISION OF GRADUATE AND UNDERGRADUATE STUDENTS

***Note:** because of the nature of the positions I held, in several cases I was only legally entitled to co-supervision.*

PhD	One ongoing (co-supervision)
Master's theses	Two completed (one as sole supervisor), one ongoing (co-supervision)
PhD project	One completed (co-supervision)
Undergraduate	One completed (research assistant; sole supervisor)

TEACHING ACTIVITIES

Spring 2014	Lecturer – Syntax and semantics (B.A.), University of Texas at Austin, USA
07/2009	Lecturer – Computational lexical semantics, ESSLLI 2009, Bordeaux, France
2008 – 2011, 2014	Lecturer – Computational linguistics, technology for translation, Linguistics (B. A, master's), U. Pompeu Fabra, Spain
2001 – 2007	Teaching assistant – Computational linguistics, technology for translation (B. A, master's), U. Pompeu Fabra, Spain

ORGANISATION OF SCIENTIFIC MEETINGS (as co-organizer)

08/2016 DisSALT: Distributional Semantics and Linguistic Theory, ESSLLI 2016, Bolzano, Italy

03/2013 Towards a formal distributional semantics, IWCS 2013, Potsdam, Germany

03/2009 *Jornada del Processament Computacional del Català* (Workshop on the Computational Processing of Catalan), Barcelona, Spain

03/2009 Nanoworkshop on Statistical Physics and Linguistics, Barcelona, Spain

07/2008 Human judgements in Computational Linguistics, COLING 2008, Manchester, UK

COMMISSIONS OF TRUST

Ongoing **Guest co-editor**, Special Issue on Formal distributional semantics, *Computational Linguistics* journal, MIT Press, USA

2016 **Area co-chair**, Semantics, ACL 2016, Berlin, Germany

2015 **Program co-chair**, *SEM 2015, Denver, USA

2015 **Local co-chair**, ESSLLI 2015, Barcelona, Spain

2015 **Project evaluator**, FONCYT (Argentina), National Science Center (Poland).

2014 – pres. **Editorial Board member**, *Linguistic Issues in Language Technologies (LiLT*) journal, CSLI, Stanford, USA (associated with the Linguistic Society of America).

2013 – pres. Information Officer, ACL SIGSEM Board

2013 **Area co-chair**, *SEM 2013, Atlanta, US

REVIEWER / PROGRAMME COMMITTEE MEMBER

Journals Artificial Intelligence, Computational Linguistics, Natural Language Engineering, ACM Transactions on Speech and Language Processing, Language Resources and Evaluation, Semantics and Pragmatics, Corpora.

Conferences 20, including several editions of ACL, EMNLP, EACL, COLING, LREC, IWCS, *SEM.

Books 2, published by John Benjamins and Springer.

Workshops 15, including several at ACL and related conferences as well as ESSLLI.

MAJOR COLLABORATIONS

Louise McNally, formal and distributional semantics / adjectives / composition / reference, Universitat Pompeu Fabra, Spain

Sebastian Padó, distributional semantics / regular polysemy, Stuttgart University, Germany

Katrin Erk, formal and distributional semantics / entailment / hypernymy, The University of Texas at Austin

Marco Baroni, distributional semantics / vision and language / reference, University of Trento, Italy

Álvaro Corral, Physics and Linguistics / Zipf's law, Universitat Autònoma de Barcelona, Spain

COMPUTATIONAL LINGUISTICS RESOURCES

More information: http://gboleda.utcompling.com/resources.

Corpora **Leader**, Wikicorpus: Freely available Wikipedia-based trilingual corpus (Catalan, Spanish, English), automatically annotated, over 750 million words.

Coordinator, CUCWEB: 166-million word Web corpus for Catalan, automatically annotated.

Tools Collaborating researcher, POS-Tagger for Old Spanish. Freely available as part of the open source suite of language analyzers FreeLing.

Collaborating researcher, CatCG: Tagger and shallow parser for Catalan.

Datasets **Leader**, four freely available (CC BY-SA) semantic datasets on adjective semantics and regular polysemy.

Collaborating researcher in a fifth dataset on the semantics of color terms.

CAREER BREAKS IN RESEARCH

29/06/2005 – 31/11/2005 Maternity leave (5 months); Barcelona, Spain

29/11/2007 – 31/05/2007 Maternity leave (6 months); Barcelona, Spain

Appendix: All on-going and submitted grants and funding of the PI (Funding ID)

<u>Mandatory information</u> *(does not count towards the page limits)*

On-going Grants

Project Title	*Funding source*	*Amount (Euros)*	*Period*	*Role of the PI*	*Relation to current ERC proposal*
LOVe: Linking Objects to Vectors in distributional semantics: A framework to anchor corpus-based meaning representations to the external world	EU – H2020-MSCA-IF-2014	180,277.20	2015-2017*	PI	This project prepares the ground for the greater challenge addressed in AMORE. Like AMORE, it concerns reference and uses distributional semantics. However, it does not develop a full model of the reference process, and it does not account for different individuated referents in a dynamic fashion: It can only address static representations of 1) public entities (“The Beatles”) o 2) single entities in one images.

*LOVe was awarded until May 2017, but it will end in September 2016 when I join the Universitat Pompeu Fabra as an assistant professor. Therefore, I will have no active projects when AMORE starts and I will be able to fully dedicate myself to it.

Section c: Early achievements track-record

RESEARCH IMPACT: SUMMARY

- Total citations to date: **530**; h-index: **12**; i-10 index: **19** (source: Google Scholar, accessed 01/11/15; self-citation counts removed).
- **Four journal publications indexed in JCR** (Journal Citation Reports of Thomson Reuters' Web of Science), **plus one** to be published in 2016 (publication 2 below).
- **17 publications in the top 10 venues in Computational Linguistics** according to Google Scholar:

1.	Meeting of the Association for Computational Linguistics (ACL)	1
2.	Conference on Empirical Methods in Natural Language Processing (EMNLP)	4
3.	North American Chapter of the Association for Computational Linguistics (NAACL)	
4.	International Conference on Language Resources and Evaluation (LREC)	6
5.	International Conference on Computational Linguistics (COLING)	3
6.	arXiv Computation and Language (cs.CL)	
7.	Computer Speech & Language	
8.	Computational Linguistics	2
9.	Conference on Computational Natural Language Learning	
10.	Language Resources and Evaluation	1

  Note: Computational Linguistics is primarily a conference-based field. Most top venues are therefore conferences (venues 1-5, 9 in the table).
- **Three** edited proceedings books, including one international conference proceedings book.
- **Two** articles in Springer volumes.

FIVE SELECTED PUBLICATIONS (self-citations manually removed)

Complete list: http://gboleda.utcompling.com/research-1/publications.

Also see Google Scholar profile.

1) G. **Boleda**; S. Schulte im Walde; T. Badia (2012). Modeling regular polysemy: A study in the semantic classification of Catalan adjectives. *Computational Linguistics*. 38 - 3, pp. 575 - 616. MIT Press. Citations: 9.

   Relevance: Article in the **top ranked** Computational Linguistics journal in the Linguistics category according to the JCR, with an impact factor of 0.940 in 2012 (journal 21 of 166 in the Linguistics category in terms of impact factor). Publication with PhD supervisors.

2) G. **Boleda**; A. Herbelot (Eds.). Special Issue on Formal Distributional Semantics. *Computational Linguistics*. In preparation, to be published in the fall of 2016. Citations: N/A.

   Relevance: As I just mentioned, *Computational Linguistics* is the **top ranked** journal of Computational Linguistics according to JCR. Special issues follow a competitive reviewing process, and only respected members of the community with a strong proposal are accepted as guest editors. Publication **without** PhD supervisors.

3) G. **Boleda**; E. M Vecchi; M. Cornudella; L. McNally (2012). First order vs. higher order modification in distributional semantics. Conference on Empirical Methods in Natural Language Processing (EMNLP), 1223 - 1233. ACL. Acceptance rate: 24%.[1] Citations: 8.

   Relevance: Article in the 2nd ranked venue in Computational Linguistics (Google Scholar). Publication **without** PhD supervisors.

4) L. McNally; G. **Boleda** (2004). Relational adjectives as properties of kinds. Olivier Bonami and Patricia Cabredo Hofherr (eds.) *Empirical Issues in Syntax and Semantics 5*. pp. 179 - 196. 2004. Citations: **110**.

   Relevance: High-impact article published as a PhD student, in a prestigious publication for selected extended contributions to the *Colloque de Syntaxe et Sémantique à Paris* (CSSP). Publication **without** PhD supervisors.

[1] See http://www.aclweb.org/aclwiki/index.php?title=Conference_acceptance_rates.

5) Corral, Á., G. **Boleda**, R. Ferrer-i-Cancho (2015). Zipf's Law for Word Frequencies: Word Forms versus Lemmas in Long Texts. PLoS ONE 10(7):doi:10.1371/journal.pone.0129031. Citations: 2.

   Relevance: Interdisciplinary research (Linguistics and Physics). This journal is in the **first quartile** of its category (Science, Multidisciplinary), with an impact factor of 3.534 according to JCR. Publication **without** PhD supervisors.

INVITED TALKS

Workshops

15/09/15 **Keynote talk**: From conceptual to referential properties with distributional semantics. *IWES: International Workshop on Embeddings and Semantics* at SEPLN 2015 (Spanish Society for Natural Language Processing conference), Alicante, Spain.

18/02/15 Distributional semantics for lexical semantics. Catalonia-Israel Symposium on Lexical Semantics and Grammatical Structure in Event Conceptualization, Jerusalem, Israel.

06/28/05 Adquisició de classes semàntiques adjectivals. III Workshop of the PhD Program in Cognitive Science and Language: "Acquisition", Barcelona, Spain.

Universities / research centers (selection from 10 in total)

17/07/13 Intensionality was only alleged: On adjective-noun composition in distributional semantics. *Guest lecture series Sonderforschungsbereich 732: Incremental specification in contex*t, **Stuttgart University**, Germany.

13/02/12 Coloring semantic spaces: Towards perceptually grounded models of word meaning. *Guest lecture series Sonderforschungsbereich 732: Incremental specification in contex*t, **Stuttgart University**, Germany.

06/09/11 Modeling regular polysemy: A study in the semantic classification of Catalan adjectives. *Linguistics Colloquiu*m, **University of Texas at Austin**, Austin, USA.

14/07/11 Modeling regular polysemy: A study in the semantic classification of Catalan adjectives. *CLIC Research Seminar* , **CIMeC, Rovereto, Italy**.

03/19/10 Computational Feedback to Linguistics: A study in the semantic classification of Catalan adjectives. *Nancy NLP Semina*r, **INRIA-Lorraine**, Nancy, France.

10/12/04 Acquiring Semantic Classes for Adjectives through Clustering. *Computational Linguistics Semina*r, **King's College**, London, UK.

11/25/04 A Quantitative Approach to the Lexical Semantics of Adjectives. *Computational Linguistics Colloquium*, **Saarland University**, Saarbrücken, Germany.

AWARDS

2015 Outstanding Reviewer Recognition, ACL 2015

2001 Extraordinary Bachelor Degree Award, U. Autònoma de Barcelona, Spain

2001 Honorable Mention, National Bachelor Degree Awards, Spanish Government

COMPETITIVE FELLOWSHIPS

2015 *Marie Skłodowska-Curie* Individual Fellowship, EU (success rate: 18.6%)

2012 *Beatriu de Pinós* post-doctoral fellowship, AGAUR, Spain (success rate: 11.7%)

2008 *Juan de la Cierva* post-doctoral fellowship, MICINN, Spain (success rate: not available)

2010 PASCAL2 European Network of Excellence, Internal Visiting Programme, EU (funding for post-doctoral visit, U. Stuttgart, Germany)

2005 PhD fellowship, *Fundación Caja Madrid*, Spain (success rate: 14.7%)

2003, 2004 Short Research Visit Programme, AGAUR, Spain (funding for two doctoral visits, Saarland U., Germany)

2001 PhD fellowship, Catalan government, Spain (success rate: 28.3%)

2000 Fellowship for the Introduction to Research, CSIC, Spain (success rate: not available)

1997 Sócrates-Erasmus scholarship, EU