DOCTORAL CANDIDATE:	Murhaf Fares
DEGREE:	Philosophiae Doctor
FACULTY:	The Faculty of Mathematics and Natural Sciences
DEPARTMENT:	Department of Informatics
AREA OF EXPERTISE:	Natural Language Processing
SUPERVISORS:	Stephan Oepen and Erik Velldal
DATE OF DISPUTATION:	23rd of August 2019
DISSERTATION TITLE:	*Noun–Noun Compound Analysis: A Holistic Perspective*

This PhD thesis sheds a new light on a widely studied linguistic construction in English called noun-noun compounds. Being a thesis in informatics and natural language processing (NLP), of course, this construction is studied from a 'computational' perspective; or more accurately, a machine learning perspective.

Studying noun-noun compounds might seem overly specialized for the wider audience. However, not only are compounds very frequently used by language speakers (which in and by itself makes studying them worthwhile), but also studying this construction servers as a means to further our efforts to computationally understand and represent human language.

Overall, the research findings and contributions of this thesis can be summarized as follows. First, this thesis presents a systematic review of past computational studies on noun-noun compounds and tries to relate them to research in theoretical linguistics. In this avenue, it becomes clear that more linguistically-informed research is likely to benefit natural language processing research in informatics. Second, the thesis questions some of the long-held views on noun-noun compound analysis in natural language processing and proposes new ways forward. Here, the thesis introduces a new dataset (i.e. computational resource) to approach noun-noun compound analysis using machine learning methods. To create this new dataset, the thesis relies on existing so-called meaning representation frameworks (that is, broadly speaking, formalisms to computationally represent sentence meaning) which not only helps situate compounds in broader meaning frameworks (unlike most past research), but also creates new links between the frameworks themselves. Third, the thesis presents s systematic and comprehensive experimentation with relatively recent machine learning methods known as transfer learning and multi-task learning. More specifically, these learning techniques are put into use to try and 'solve' the task of interpreting noun-noun compounds (that is, to 'train' machine learning models to predict the semantic relation between the nouns in a given compound). The difficulty of this task is inherently dependent on how we choose to represent the semantic relations between nouns in compounds. In the case of this thesis, this interpretation task falls on the more difficult side and therefore the experiments presented help determine the utility of transfer and multi-task learning techniques in hard-to-learn tasks.