

Chapter 3

Discourse Analysis and ANLP

Alexandra Kent

Loughborough University, UK

Philip M. McCarthy

The University of Memphis, USA

ABSTRACT

The goal of this chapter is to outline a (primarily) qualitative and (secondarily) quantitative approach to the analysis of discourse. Discourse Analysis thrives on the variation and inconsistencies in our everyday language. Rather than focusing on what is said and seeking to reduce and homogenise accounts to find a central meaning, discourse analysis is interested in the consequences of “saying it that particular way at that particular time.” Put another way, it is interested in “what was said that didn’t have to be, and why?” and “what wasn’t said that could have been, and why not?” The chapter outlines the basic theoretical assumptions that underpin the many different methodological approaches within Discourse Analysis. It then considers these approaches in terms of the major themes of their research, the ongoing and future directions for study, and the scope of contribution to scientific knowledge that discourse analytic research can make. At the beginning and end of the chapter, we attempt to outline a role for Applied Natural Language Processing (ANLP) in Discourse Analysis. We discuss possible reasons for a lack of computational tools and techniques in traditional Discourse Analysis but we also offer suggestions as to the application of computational resources so that researchers in both disciplines might have an avenue of interest that assists their work, without directing it.

INTRODUCTION

Discourse Analysis is a big tent. Various people, who each would describe themselves as discourse analysts, employ a wide range of different methods and may make very different sets of assump-

tions (Phillips & Jorgensen, 2002). That having been said, and at a very basic level, we can assume that Discourse Analysis is the analysis of data involving words (either spoken or written). Generally, this analysis is part of talk in a social interaction (Howitt, 2010), although sometimes the analysis can reach the level of an entire text (Trappes-Lomax, & Hughs, 2004); at which

DOI: 10.4018/978-1-60960-741-8.ch003

point, Discourse Analysis might be said to merge somewhat with field of Text Analysis. Discourse Analysis is most often a qualitative methodology, although post-positivistic discourse analysts often accept the assistance of (low level) computational and statistical approaches so as to gain a broader perspective of the subject of interest (Denzin & Lincoln, 2005; Gee, 2005). In such a big tent, it is perhaps not surprising that few researchers agree on how Discourse Analysis should be conducted. For example, some researchers advocate that only naturally occurring data should be used, whereas others feel discourse analysis can be fruitfully applied to researcher generated data such as interviews or focus groups (McKinlay & McVittie, 2008). To this end, Howitt (2010) offers a neat map of several different varieties of Discourse Analysis, showing how they relate to each other (2010: 217).

On the face of it, the field of Discourse Analysis would seem like an obvious complement to the theme of this book: Applied Natural Language Processing (ANLP). After all, the word *discourse* in Discourse Analysis clearly relates closely to the words *natural language* in ANLP; and the word *analysis* in Discourse Analysis relates closely to the word *processing* in ANLP. However, the reality is that the field of Discourse Analysis lies far from more prototypical ANLP pursuits such as Natural Language Understanding and Processing (see Rus, this volume), Corpus Linguistics (see Hall this volume), Data Mining, (see Atzmueller this volume) and so forth. This distance between Discourse Analysis and the more prototypical fields results from the focus of studies that are of typical interest to Discourse Analysts. These interests do not easily or obviously lend themselves to computational approaches. Such an issue may reasonably lead readers to ask why a chapter on Discourse Analysis is even necessary in this book. Our response to such a question is to argue that while computationally addressing areas of interest to Discourse Analysis are, undoubtedly, challenging, they are, nevertheless, opportunities

for resourceful researchers who wish to apply developments in textual technology. As such, we view studies in Discourse Analysis as a largely untapped reservoir of ANLP research.

In this chapter, we begin with a brief discussion as to why computational (quantitative) approaches have been relatively rare in Discourse Analysis. We then discuss the basic assumptions of Discourse Analysis before outlining some of the major fields of interest in Discourse Analysis. We end with our thoughts on where and how ANLP research might be employed so as to enhance Discourse Analysis research.

DISCOURSE ANALYSIS AND ANLP (OR LACK THEREOF)

Discourse Analysis does not have a strong tradition in ANLP. Three major issues help to explain this observation. First, although studies in Discourse Analysis can focus on spoken or written text, it is probably fair to argue that the majority of studies (and the best known studies) are on the computationally less favourable spoken form. Second, Discourse Analysts are frequently interested in what was not said, that could have been said. From a computational point of view, it is difficult enough to evaluate with any accuracy what the text actually is, let alone what it could have been. The third point relates closely to the second. For Discourse Analysts, it is not just *what* was not said, but *why* it was not said, and what might this “choice” tell us about our world, our societies, and ourselves. Clearly, such interests are predominantly qualitative pursuits that seem far from the quantifications that are more popular with computational approaches.

TALK AS TOPIC

Discourse Analysts are interested in what talk does, not what it is about (Wood & Kroger, 2000). This

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/discourse-analysis-anlp/61041

Related Content

Sentiment Classification: Facebook' Statuses Mining in the "Arabic Spring" Era

Jalel Akaichi (2015). *Modern Computational Models of Semantic Discovery in Natural Language* (pp. 1-26).

www.irma-international.org/chapter/sentiment-classification/133873

Shape Grammar Approaches in Architecture

(2020). *Grammatical and Syntactical Approaches in Architecture: Emerging Research and Opportunities* (pp. 19-55).

www.irma-international.org/chapter/shape-grammar-approaches-in-architecture/245859

A Customizable Language Learning Support System Using Ontology-Driven Engine

Jingyun Wang, Takahiko Mendoriand Juan Xiong (2014). *Computational Linguistics: Concepts, Methodologies, Tools, and Applications* (pp. 724-739).

www.irma-international.org/chapter/a-customizable-language-learning-support-system-using-ontology-driven-engine/108748

An Overview and Technological Background of Semantic Technologies

Reinaldo Padilha França, Ana Carolina Borges Monteiro, Rangel Arthurand Yuzo Iano (2021). *Advanced Concepts, Methods, and Applications in Semantic Computing* (pp. 1-21).

www.irma-international.org/chapter/an-overview-and-technological-background-of-semantic-technologies/271118

Co-Designing Participatory Tools for a New Age: A Proposal for Combining Collective and Artificial Intelligences

José Luis Fernández-Martínez, Maite López-Sánchez, Juan Antonio Rodríguez Aguilar, Dionisio Sánchez Rubioand Berenice Zambrano Nemegyei (2020). *Natural Language Processing: Concepts, Methodologies, Tools, and Applications* (pp. 860-877).

www.irma-international.org/chapter/co-designing-participatory-tools-for-a-new-age/239970