Chapter 31
Evaluation of Narrative and Expository Text Summaries Using Latent Semantic Analysis

René Venegas
Pontificia Universidad Católica de Valparaíso, Chile

ABSTRACT

In this chapter I approach three automatic methods for the evaluation of summaries from narrative and expository texts in Spanish. The task consisted of correlating the evaluation made by three raters for 373 summaries with results provided by latent semantic analysis. Scores assigned by latent semantic analysis were obtained by means of the following three methods: 1) Comparison of summaries with the source text, 2) Comparison of summaries with a summary approved by consensus, and 3) Comparison of summaries with three summaries constructed by three language teachers. The most relevant results are a) a high positive correlation between the evaluation made by the raters (r= 0.642); b) a high positive correlation between the computer methods (r= 0.810); and c) a moderate-high positive correlation between the evaluations of raters and the second and third LSA methods (r= 0.585 and 0,604), in summaries from narrative texts. Both methods did not differ significantly in statistical terms from the correlation among raters when the texts evaluated were predominantly narrative. These results allow us to assert that at least two holistic LSA-based methods are useful for assessing reading comprehension of narrative texts written in Spanish.

INTRODUCTION

In Latin America, comprehension skills associated with reading and writing are not effective enough for students to understand and produce texts that will allow them to perform well in today's society (Peronard, 1989; Peronard, Gómez,

Parodi & Núñez, 1998; PISA, 2007; Ibañez, 2008). There is general consensus between researchers studying reading comprehension as to the need of evaluating individuals psycholinguistic processes, highlighting several options for evaluation, such as answers to literal and inferential questions (both local and global), development of conceptual and/or mental maps, formulation of open-ended questions, paraphrasing, and summary writing.

I am specifically interested in further inquiry regarding one of these options, which is the summary as an evaluation technique. The use of this technique is theoretically supported by the proposals of van Dijk (1978). From this psycholinguistic perspective, we attempt to formalize processes by the reader and subsequently incorporate them into memory. I first propose the so-called macrorules (van Dijk, 1978) for this purpose, followed by macrostrategies (van Dijk & Kintsch, 1983) to attempt an explanation as to why what person remembers and verbalizes after reading a relatively long text does not include all the ideas originally expressed in such texts. Good comprehenders apply rules and strategies, eliminating propositions they believe to be irrelevant, and reprocessing others in order to build their own version of the text. However, this evaluation technique features some problems, mainly regarding human variables such as the cognitive burden of the evaluator, parallel attention to formal elements of written production (spelling, handwriting), subjective aspects that may intervene, systemacity in the application of evaluation criteria, consensus between multiple evaluators, and the extensive amount of time required for evaluation.

These problems and readers interest in automatically capturing core text information have encouraged the study of summarizing and summary evaluation from a computational perspective. Automatic summary construction and automatic summary evaluation are problems that have been discussed since the mid-60s, although computer systems reliable enough to perform both tasks have not yet been found. However, advances based on computer linguistics, natural language processing, and the development of several information recovery techniques lead us to think that we are closer to improving the generation and evaluation of written summaries (Mani & Maybury, 2001).

Techniques for building and automatically evaluating text summaries are generally classified into two categories: linguistic and statistical. Linguistic techniques use knowledge regarding syntax, semantics or the use of language, while statistical techniques operate by computing values for words and phrases found in the text, using statistical techniques such as frequencies, n-grams, and co-occurrences.

There are several studies focusing on comprehension and the use of computer techniques in order to represent and evaluate the comprehension process. Kintsch (1998, 2000, 2001, and 2002) introduces the possibility of using latent semantic analysis (LSA) for extracting lexico-semantic similarities from the texts and accessing text propositions by means of statistical-mathematical training to simulate the comprehension process.

Theoretical and empirical studies concerning LSA have been extensively conducted in different languages, but most often in English (e.g., Landauer & Dumais, 1997; Landauer, Laham, Rehder & Schreiner, 1997; Rehder, Schreiner, Wolfe, Laham, Landauer, & Kintsch, 1998; Kintsch, Steinhart, Stahl, LSA Research Group, Matthews, & Lamb, 2000; Landauer & Psotka, 2000; Landauer, Laham, Foltz, 2003). These studies have focused on evidencing that vectorial models, and specifically LSA, can account for meaning processing conducted by human beings and are effective in the automated evaluation of essays and summaries. However, studies considering LSA in the field of summary evaluation in Spanish are still scarce (Pérez, Alfonseca, Rodríguez, Gliozzo, Strapparava & Magnini, 2005; León, Escudero, Cañas & Salmerón, 2005; Venegas 2007, 2009a, b). Therefore, this research aims to identify an automatic summary evaluation method for texts written in Spanish that are predominantly narrative and expository, using latent semantic analysis.

SUMMARY AS A TECHNIQUE FOR EVALUATING THE UNDERSTANDING OF WRITTEN DISCOURSE

Two types of semantic structures can be recognized in a text: microstructures and macrostructures

Related Content

The Role of Textual Graph Patterns in Discovering Event Causality

Bryan Rink, Cosmin Adrian Bejanand Sanda Harabagiu (2012). *Applied Natural Language Processing: Identification, Investigation and Resolution (pp. 334-350).*

www.irma-international.org/chapter/role-textual-graph-patterns-discovering/61057

The Language of Glenn Murcutt's Domestic Architecture

 (2020). *Grammatical and Syntactical Approaches in Architecture: Emerging Research and Opportunities (pp. 215-262).*

www.irma-international.org/chapter/the-language-of-glenn-murcutts-domestic-architecture/245865

Neural Network Model for Semantic Analysis of Sanskrit Text

Smita Selot, Neeta Tripathiand A. S. Zadgaonkar (2020). *Natural Language Processing: Concepts, Methodologies, Tools, and Applications (pp. 1011-1025).*

www.irma-international.org/chapter/neural-network-model-for-semantic-analysis-of-sanskrit-text/239977

Learning Languages via Social Networking Sites

Billy Brick (2014). *Computational Linguistics: Concepts, Methodologies, Tools, and Applications (pp. 763-778).*

www.irma-international.org/chapter/learning-languages-via-social-networking-sites/108750

Towards a Bio-Inspired Theoretical Linguistics to Model Man-Machine Communication

Gemma Bel-Enguixand M. Dolores Jiménez-López (2014). *Computational Linguistics: Concepts, Methodologies, Tools, and Applications (pp. 1422-1437).*

www.irma-international.org/chapter/towards-a-bio-inspired-theoretical-linguistics-to-model-man-machine-communication/108785