

The Legacy of János S. Petőfi

The Legacy of János S. Petőfi:

*Text Linguistics, Literary
Theory and Semiotics*

Edited by

Margarita Borreguero Zuloaga
and Luciano Vitacolonna

Cambridge
Scholars
Publishing



The Legacy of János S. Petőfi:
Text Linguistics, Literary Theory and Semiotics

Edited by Margarita Borreguero Zuloaga and Luciano Vitacolonna

This book first published 2019

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2019 by Margarita Borreguero Zuloaga, Luciano Vitacolonna
and contributors

All rights for this book reserved. No part of this book may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording or otherwise, without
the prior permission of the copyright owner.

ISBN (10): 1-5275-2310-1

ISBN (13): 978-1-5275-2310-4

CONTENTS

Introduction

The Lasting Legacy of János S. Petőfi	viii
Margarita Borreguero	

In Memory of János S. Petőfi	xx
Luciano Vitacolonna	

Linguistics

On the Text Organizing Status of Key Words in News Discourse: An Empirical Study in Terms of János S. Petőfi's Legacy	2
József Andor	

From Text Grammar to Macropragmatics via Coherence	24
Klaus Hölker	

Petőfi's View of Language and Current Cognitive Linguistics	34
Enrique Bernárdez	

Corpus-based Text Linguistics	53
Manuel Barbera, Carla Marelllo	

On a Computational Semiotic Textology	69
Borja Navarro Colorado	

Literary Theory and Rhetoric

The Pragmatics in János S. Petőfi's Text Theory and the Cultural Rhetoric: The Extensional-Semantic Code and the Literature of the Spanish Golden Age	92
Tomás Albaladejo	

János S. Petőfi's Linguistic and Textual Theory and the Recovery of the Historical Thinking about Rhetoric	110
Francisco Chico Rico	

Philology within the Framework of Semiotic Textology: János S. Petőfi's Scientific Legacy II	132
Mauro Giuffrè	

Semiotics

Semiotic-Textological interpretation of Communication in Art Museum	168
Giuliana Pascucci	

Appendix

On Basic, Disciplinary Issues of Textology and the Status of Semiotic Textology: An Electronic Interview with János Sándor Petőfi.....	182
József Andor	

INTRODUCTION

THE LASTING LEGACY OF JÁNOS S. PETŐFI

MARGARITA BORREGUERO ZULOAGA

János S. Petőfi (1934-2013) was a Hungarian linguist and semiotician who developed his research and teaching activities at the Academy of Sciences in Budapest and at the Universities of Bielefeld (Germany) and Macerata (Italy), among others, until his retirement in 2009. He had a mathematical background and was also a graduate of German Studies at the University of Debrecen (1961). After leaving Hungary for political reasons and obtaining his PhD at the University of Umeå (Sweden) in 1971, thanks to a grant for political refugees, he moved to Germany and became part of the initial group of researchers that founded a new linguistic discipline, Text Linguistics, at the University of Konstanz (Germany) under the leadership of Peter Hartmann in the late '60's and early '70's.

In the early stage of his research career in Western Europe, he developed a very powerful heuristic formal model inspired by the theoretical framework of generative grammar to explain the syntactic and semantic links between sentences in texts; this model was later improved with an extensional semantic component and named Text Structure World Structure Theory (TeSWeST in its German abbreviation). In the '90s it turned into a semiotic model when Petőfi began to deal with texts in which verbal, musical, kinesic and graphic elements were intertwined. The theory was renamed Semiotic Textology and was applied to the analysis of songs, comics, illustrated books, TV-news, Greek inscriptions, graphic poetry and a wide range of texts by him and his collaborators and disciples.

His amazing scientific production (over 400 titles in Hungarian, English, German, Italian, Spanish, French and Japanese) had a strong impact on researchers dealing with Linguistics, Literary Theory, Philosophy of Language, Computer Linguistics and Semiotics all over Europe and beyond. He set up a network of colleagues and disciples—who often visited him in Germany and Italy as well as inviting him to teach or take part in conferences and seminars—and tutored the work of many PhD students in different countries. In the last years of his teaching career, after his retirement and return home to Hungary, he was mainly interested in restoring links with Hungarian universities and academia and, together

with his last wife Zsuzsa Benkes, devoted much of his research looking for a way to present his ideas about text in the teaching of the Hungarian language at school.

In these introductory pages, we do not intend to highlight Petőfi's main theoretical achievements, nor do we wish to display his many activities as a scientific promoter, which include the editorship of the two collections *Papiere in Textlinguistik* (Buske) and *Research in Text Theory* (De Gruyter) and being editor-in-chief of the journal *Text*. But rather, we will refer to many of his most important contributions to the above-mentioned scientific fields when presenting the papers gathered in this volume (see below).

His sharp intelligence, enthusiasm for research and kindness moved many students and young researchers to follow in his footsteps in the field of Semiotics and Text Linguistics. When he passed away on February 10th, 2013, most of us were not able to reach Budapest on time to attend his funeral. This deepened our sorrow and unbearable pain for his loss which we are trying to alleviate by continuing his work, organising international conferences *in memoriam* and publishing new research around his legacy in collective volumes, following the path set by the several *Festschriften* published to mark his 80th birthday.¹

The First International Workshop *János S. Petőfi in memoriam* took place on April 23rd, 2015, at the Faculty of Philology of the Complutense University of Madrid. The date of the workshop was not chosen by chance, as it would have been Petőfi's 82nd birthday. The organizers invited a group of colleagues, friends and disciples of János S. Petőfi from Hungary, Germany, Italy and Spain. Each of them had contributed to or applied Petőfi's theoretical framework at several stages of its development, besides maintaining a close relationship with him either as colleague or as a student (or both). This workshop, in an intimate and fraternal atmosphere, gave us the opportunity both to remember a master and to confront our contributions accomplished with inspiration from Petőfi's thought and work. As such, the proceedings of the workshop

¹ Hölker, Klaus; Marelllo, Carla (eds.), *Dimensionen der Analyse von Texten und Diskursen. Festschrift für Janos Sandor Petőfi zum achtzigsten Geburtstag*, London-Berlin-Münster, Lit Verlag, 2011; M. Giuffrè (ed.), *Studies in Semiotic Textology in Honour of János S. Petőfi. Supplement to Sprachtheorie und germanistische Linguistik* (2011).

constitute not only a sign of gratitude towards a great mentor, but also proof of the continuity and vitality of his teachings.²

The articles have been grouped into three different parts. The first of which is devoted to Linguistics (Text and Computational Linguistics) and contains the contributions by József Andor (Pest), Klaus Hölker (Hannover), Enrique Bernárdez (Complutense, Madrid), Manuel Barbera and Carla Marello (Turin), and Borja Navarro (Alicante). Both Andor and Hölker's papers deal with the topic of textual coherence, which play a key role in Petőfi's text concept, while Bernárdez, Barbera-Marello and Navarro explore the connection of his textual and semiotic theories with other linguistic disciplines, Cognitive Linguistics, Corpus Linguistics and Computer Linguistics, respectively.

József Andor, in his paper "On the text organizing status of key words in news discourse: An empirical study in terms of János S. Petőfi's legacy", describes the scientific context in which Petőfi's text theory emerged and how most of the scholars in this field, following Halliday and Hasan's pioneer work, gave priority to the textual cohesive mechanisms over the concept of textual coherence. Petőfi was conscious of the key role played by coherence in the interpretation of texts, which is not based on cohesion and connectivity, but on world knowledge conceived of as conceptual schemata (the so-called frames and scripts, among others). What Andor proposes in his contribution is to show how Petőfi's wide concept of coherence can be effectively applied to text analysis. He does so by analysing the contribution of the lexical component at the macrostructure level, and precisely the coherence capacity of key words in journalistic texts. "Due to the linguistic representation of sense, words have the capacity to evoke and instigate frames and scenes from the conceptual lexicon. As such, their sense relations are mentally organized in sets of conceptual hierarchies. Items which have a central role in representing frames in the mind and which have sets of bodies of lexis under their control, conceptually related to them under the dominance of a frame, have a key status in the textual representation of their underlying frame. Such words are key words." (Andor, this volume). Andor's concept of key words is not based on their frequency in language or in a particular text type, but in their capacity to evoke conceptual frames.

The choice of journalistic texts as a field of analysis is due to the fact that key words have a more organised distribution in this kind of texts (around the heading and first paragraphs) and contribute in a very effective

² Good proof of the impact of János S. Petőfi's work and personality is his *honoris causa* at the University of Turin and of Pest and the János S. Petőfi Prize created for graduate students at the Faculty of Philosophy of Macerata University.

way to the interpretation process. Andor compares the results of a frequency test on two newspaper articles dealing with the same subjects to the results of a group of native speakers selecting the most relevant words for the understanding of those texts and concludes that only in the second case are the selected words significant for text coherence as far as their function is to sign frame-based representations, i.e. “conceptualization, setting of the scene, evoking and activation of the relevant background knowledge, common ground between the journalist and the readers” (Andor, this volume). He closes his paper with the wish that research on text coherence will continue along the path envisioned by Petőfi.

Klaus Hölker’s paper “From Text Grammar to Macropragmatics via Coherence” presents the historical context in which Petőfi’s interest on textual coherence was aroused, a context he himself was part of at the beginning of his career at the University of Bielefeld. Coherence was a latter concept in the history of Text Linguistics since this discipline emerged from generative syntax and in its first years of life was conceived of as an extension of sentence syntax and therefore mainly focused on formal aspects, such as cohesive and connexity devices between utterances. But Petőfi, as early as in 1971, proposed to overtake the conception of a co-textual text theory and built a contextual text theory to take into account “encyclopaedic knowledge and the poetic, aesthetic, sociological and ideological dimensions of texts” (Hölker, this volume), as well as the role of the communication situation and the speech act theory.

It was only in the ‘80’s that text linguists realized that cohesion and connexity were not sufficient conditions to speak about textuality, and that world knowledge was crucial in text interpretation, as the failure of automatic translation had already shown. The problem was then how to integrate world knowledge in text theory, a topic around which Petőfi organised several workshops. For Petőfi it was clear that coherence depended on the interpreter and was not a characteristic of the text itself (Hölker and Bernárdez, this volume), but he tried to further refine this concept and came to the conclusion that “a text is coherent if and only if the interpreter of a text succeeds in constructing a mental image of the extralinguistic reality described by the text in such a way that all of the states of affairs described appear as interconnected by relevant relations” (Hölker, this volume). According to Hölker, this concept of coherence led Petőfi to redefine the semantic component of his last textual theory, Semiotic Textology, and to give a more prominent role to extensional semantics in the interpretation of the lexical unit (the *sensus*) and to the correspondence between the lexical unit and the mental representation of the extralinguistic reality (the *relatum-imago*).

The inspiring character of this model to contemporary Linguistics is also to be seen in Enrique Bernárdez's paper "Petőfi's view of language and current Cognitive Linguistics". Notwithstanding his ideas about the above-mentioned relation between world knowledge and textuality, Petőfi's work has been largely ignored in Cognitive Linguistics. As Bernárdez recalls, he travelled to the US, met Lakoff and presented his textual model at the Berkeley Linguistic Department in 1976, but it apparently had little to no impact in the gestation of this new linguistic view. What Bernárdez proposes in this paper is that Petőfi's theory has some strong points that could have bolstered some of Cognitive Linguistics' present weaknesses if more attention had been paid to his work. One of the drawbacks of Cognitive Linguistics is, for example, that sentences and words continue to be considered the main linguistic units and the cognitive dimension of texts has been almost unexplored, particularly by American scholars, less so by their European counterparts.

As of his very first works, Petőfi highlighted the interactive nature of text, which is the result of a co-construction between participants at the communicative situation, and this conceptualization has been largely accepted in linguistics today, including Cognitive Linguistics. But, according to Bernárdez, Petőfi's methodological views have not been taken into account by most cognitive linguists, such as his focus on the necessity of using canonical formal metalanguage for linguistic description and of basing the analysis in real texts (and not in pre-constructed examples), which through (dis)confirmation of theoretical claims will allow the theory to move forward. Regarding more specific cognitive core ideas, the correspondence between text and the mental representation of external reality which is agreed upon by the interlocutors solves the debate between objectivism and relativism. "What matters is not so much the 'reality', the 'state-of-affairs' underlying linguistic utterances, but the mental images of that reality agreed upon by the participants. JSP's view of the reality underlying language is more authentically cognitive than the one encountered in many a CL analysis" (Bernárdez, this volume). His semiotic conception of text has also been ignored except in the case of Brandt's Cognitive Semiotics.

As a conclusion, "Petőfi's view is undoubtedly cognitive: the centre of what can be termed "textual process" is the participants dealing with mental images in real or assumed—but always cognitively constructed, interpreted—situations of communication. [...] The main difference, I dare say, is that CL has frequently been little systematic in the application of its own epistemological concepts; in contrast, JPS's models show complete systematicity and precision" (Bernárdez, this volume). Behind these

observations lies the sharp criticism on how different schools in linguistics refused to accept and integrate ideas from scholars belonging to other schools or groups into their own theories and analyses (they even refused to acquaint themselves with and acknowledge those ideas), what Bernárdez called the *abyssal thinking* and the *invertebrate* nature of Linguistics as scientific discipline.

Manuel Barbera and Carla Marello, in “Corpus-based Text Linguistics”, explore the relationship between Corpus linguistics and Text linguistics, a relationship Petőfi was rather sceptical about. His scepticism has a plain explanation if we think about Petőfi’s generative grammar background and remember how critical Chomsky was towards corpus linguistics. Even when corpus linguistics was in the position to produce more accurately collected corpora that could be useful for text linguistics research, Petőfi considered that reliability of corpora was similar to that of introspection.

However, in order to carry out solid corpus-based textual research, the Hungarian linguist deemed it necessary to have video and audio recordings in spoken corpora so as not to miss the important meaning conveyed by paralinguistic and extralinguistic elements. This kind of corpora are not available at large yet, but scholars agree about the importance of multimodal corpora for an accurate analysis of human communication. In fact, some achievements such as semiotic mark up and multimodal query systems have been elaborated following the path advocated by Petőfi. While multimodal corpora were taking their first steps, Petőfi had already transformed his mainly verbal textual model (TeSWeST) into a semiotic interdisciplinary one, Semiotic Textology, which remains to this day the most powerful theoretical proposal for the analysis of semiotic texts, i.e. all kinds of texts since there are no monomedial ones.

This section closes with Borja Navarro’s paper on the contribution of Semiotic Textology to Computational Linguistics. During his stay at the Academy of Sciences in Budapest, Petőfi had a strong interest in the development of Computational Linguistics and wrote several papers on this topic (mostly in Hungarian) putting into play both his mathematical and linguistic backgrounds. In his “On a computational Semiotic Textology”, Navarro points out the connections between the semantic component of the Semiotic Textology and the new Vector Space Model for Semantics, a new computational approach to language interpretation and understanding. Both models are interested in representing not only the linguistic component of texts, but also other semiotic systems such as music and image.

According to Navarro (this volume), “Semiotic Textology could shed light on two challenges now at stake in Computational Linguistics: the compositional-distributional semantic models and their extension to a unified multimedia semantic representation”. One aspect in which Semiotic Textology could be useful to the Vector Space Model is by offering a theoretical framework to the concept of distributional meaning based on vector spaces (i.e. possible and actual context of occurrence of a lexical unit). One drawback of the Vector Space Model is the conceptualization of lexical meaning, intentionally conceived and reduced to the different contexts in which the word appears. On the contrary, the lexical meaning in Petőfi’s model is considered in all its complexity: On the one hand, Petőfi distinguishes between the meaning attributed to an element in a semiotic system (*sensus sistemicus*) and the meaning acquired in a communicative situation (*sensus contextualis*), on the other hand, he considers three different dimensions (relational, inferential and configurational). Navarro claims that the inferential dimension of meaning in Petőfi’s model is similar to the distributional meaning in Vector Space Model. “During the interpretation, the interpreter infers the meaning of a word according to the similarity of its context compared to all the previous contexts in which this word has appeared. From a vectorial point of view, this inference is based on the similarity between the new contextual vector of the word and the previous vectors in the semantic space. The interpretation is coherent if the new context and the previous contexts are in some way similar” (Navarro, this volume)

Both models also converge regarding the process of interpretation. Semiotic Textology modelised this process through different types of cognitive bases, containing knowledge about the world, hypotheses and preferences. In the Vector Space Model, the knowledge about the meaning of a lexical unit comes from the different contexts in which the unit appears; all of these contexts have to be stored somehow in the interpreter’s cognition. According to Petőfi, the place to be stored is the General Typological Basis (B_{Tp}). In this way, the Petőfian model offers again a framework to explain how interpretative processes are based on similarity.

The last issue tackled by Navarro is how the Semiotic Textology framework could contribute to the development of Multimedia Distributional Vector-based Semantic Representations. Petőfi’s model distinguishes three kinds of meaning linked to different semiotic system: the *dictum* (conceptual verbal meaning), the *apperceptum* (conceptual non-verbal meaning) and the *evocatum* (non-conceptual meaning). The Vector Space Model focuses only on the *dictum* and Navarro discusses its

present limitations (and attempts) to deal with non-verbal meaning (computational models have to establish a unit compatible with a verbal textual unit) and with the mental representations of reality. On the other hand, Vector Space Models could provide Semiotic Textology with formal language to represent any kind of media. Hence, Navarro advocates for an integration of both models.

The second part of the book is devoted to the impact of Petőfi's textual models in the disciplines that have traditionally developed around the analysis of text production and reception: Rhetoric, Poetics and Literary Theory. It contains the contributions of Tomás Albaladejo (Universidad Autónoma de Madrid), Francisco Chico Rico (Alicante) and Mauro Giuffrè (Palermo). The first two authors devoted several studies to the implementation of a pragmatic component and a metatheoretical component to the last version of the TeSWeST at the beginning of the '80's. In their contributions, they remind us of how literary texts were present in Petőfi's research as of his very first papers and how much literary analysis owes to Text Linguistics.

Albaladejo, in his "The Pragmatics in János S. Petőfi's Text Theory and the Cultural Rhetoric. The Extensional-Semantic Code and the Literature of the Spanish Golden Age", reminds us how many of the thoughts of classical Rhetoric regarding the production and interpretation of texts were integrated in Text Linguistics and how Petőfi himself considered them in the outline of his first textual models. Therefore, he advocates the validity of Petőfi's textual model for the new rhetoric studies, such as Cultural Rhetoric which aims at studying literature, discourse and culture "considering their perlocutionary components, mainly those related to persuading and convincing [...] [and] is oriented to the analysis and the explanation of the function of culture in the pragmatically oriented constitution of literary works and other kinds of discourses" (Albaladejo, this volume).

Albaladejo pays particular attention to the tripartite structure of what is probably Petőfi's most famous text model, TeSWeST, and reflects upon the construction of such a model taking into account the three disciplines that integrate Semiotics according to Morris: Syntax, Semantics and Pragmatics. But while the syntactic and the semantic components of the model are clearly differentiated, the pragmatic component never had an independent representation within it, because—quoting Petőfi—"no independent pragmatic component can be separated within a theory of language" (see Albaladejo and Hölker, this volume). Both syntax and semantics have a pragmatic dimension and what Albaladejo has proposed in many of his writings is to make explicit this pragmatic dimension of the

Textual and World components in the TeSWeST. Moreover, since the theory distinguishes at least two types of receiver, the plain receiver and the analyst, Albaladejo proposes an Extended TeSWeST “to separate a meta-theoretical level from the theoretical level within the text model [...]”. The representation component has been built and works as a construction reflected from the general part of the model” (Albaladejo, this volume). The combination of this two-fold enlargement of the TeSWeST led to the E TeSWeST II, which is considered a suitable model for contemporary literary and rhetorical studies.

Albaladejo further reflects on the importance of sharing not only a linguistic code but also a word model with the producer of the text to achieve a successful textual interpretation. In fact, “[t]he extensional-semantic elements constituting referents also work as a code. It is the case of historical, literary, mythological, religious, philosophical and artistic elements that are part of referents as cultural topics” (Albaladejo, this volume) and that is the reason why Cultural Rhetoric is so interested in word models and could gain some insight from textual models with extensional-semantic and pragmatic components such as the E TeSWeST.

The relationship between the TeSWeST and Rhetoric is also the topic of Chico’s paper “János S. Petőfi’s linguistic and textual theory and the recovery of the historical thinking about Rhetoric”. The guiding principle of his contribution is the fruitful cooperation between classical Rhetoric and Poetic, on the one hand, and Text Linguistics and Linguistic Poetics, on the other, to build a General Rhetoric as has been advocated for by theoreticians such as García Berrio and Albaladejo. In particular, Chico is interested in showing how a redefinition of the rhetoric operation *elocutio* could benefit from the TeSWeST theoretical framework, especially from the above-mentioned extended version E TeSWeST II, and how simultaneously the study of rhetorical *elocutio* could complement E TeSWeST II giving rise to E TeSWeST III.

He focuses on how the semantic-intentional information is transformed into a linear text structure through a mapping component and considers that the original transformation operation in the TeSWeST has to be developed in order to distinguish between two different operations (*dispositio* and *elocutio*). The result of the *dispositio* operation in narrative texts, as pointed out by Albaladejo in previous works, is the chronological and causal order of events, the macrosyntax; the result of the *elocutio* operation is the linear text structure. Chico explains that *elocutio* operates directly on the lexicon component, whose mapping component produces the manifestation macrosyntactic level, and thus enriches the original TeSWeST which is then renamed E TeSWeST III. According to this

author, the rhetoric operations can be redefined in correspondence to the E TeSWeST III model. “This redefinition implies a definitive abandonment of successive and compartmentalised thought about the operations which form the rhetorical system as a descriptive and explanatory model for text construction and communication” (Chico, this volume). Finally, he also advocates for the introduction of a previous noetic operation, called *intellectio*, which is responsible for the adequation of the rhetoric operations to the communicative situation and matches the pragmatic component of the E TeSWeST III. This paper effectively shows how the combination of Rhetoric and Petőfi’s original and extended models produces more accurate descriptions of the text production process.

In the last paper in this section, Giuffrè deals with the interesting question of the interdisciplinary context of Semiotic Textology, i.e. the relationship to other disciplines such as Linguistics, Semiotic, Communication Theory, Philosophy, etc. Petőfi did not offer a fixed list of disciplines but rather considered that the number and type of disciplines coming into play in the interpretation process depends on the type and nature of the text subject to interpretation: The disciplines relate to the conventions, beliefs and world knowledge of a particular text and to the semiotic systems at play. Petőfi devoted several papers to the possible configurations of this interdisciplinary network around Semiotic Textology, as Giuffrè shows in his contribution, proposing several classifications of the disciplines and relationships among them.

Among all these disciplines, Giuffrè focuses on Philology and its relationship with Semiotic Textology as declared in the title of his paper: “Philology within the framework of Semiotic Textology”. He establishes a four-phase interpretive process called Natural Text Processing (NTP)—in analogy to computer-mediated Natural Language Processing (NLP) which he extensively presents—: describing, analysing, explaining and evaluating a text. While the former two are accomplished by Text Linguistics, the third is attributed to Hermeneutics and the latter to Literary Criticism. The reason why scholars working in these disciplines might be interested in Semiotic Textology is that “Petőfi’s Semiotic Textology makes this interdisciplinary NTP module [...] reticular, flexible and interdependent, but not sequential because all phases do not work separately as a fixed procedure but inter-react reciprocally” (Giuffrè, this volume). The paper dwells on the different schools, authors and conceptions that have shaped Philology as a scientific discipline and accepts Scibetta’s distinction between a *stricto sensu* Philology—which deals with the formal aspects of the text and is previous to NTP—and a

lato sensu Philology—which deals with its historical context and is central to the phases of analysis and explanation—.

In the framework of the Semiotic Textology, Philology *stricto sensu* or Textual Criticism can be considered a Linguistic discipline (L-discipline), as well as a Formal and an Empirical Methodology; while Philology *lato sensu* is considered a Specific discipline (S-discipline) providing the anthropological-cultural knowledge required for text interpretation as well as what Petőfi calls an X-discipline, a specific discipline for a particular text, as it “organises systems and patterns of conviction, belief, ideology, religion and systems of knowledge of the textual world contained in each analysed text” (Giuffrè, this volume).

The last part of the book contains just one article on Semiotics by Giuliana Pascucci (Macerata), “A semiotic-textological approach to communication in art museums”. Pascucci’s contribution represents the kind of research Petőfi was more interested in during his last years at the University of Macerata: Communication processes and complex signs involving multiple semiotic systems that interrelate to create a multimedial text. A worthy example of this kind of communication process are those which take place at art museums. After a very noteworthy introduction about the role of museums in contemporary society, Pascucci deems Petőfi’s Semiotic Textology a suitable theoretical framework to rethink communication in the context of art museums as a communicative process which has the visitor at its very centre. On the one hand, she accomplishes the task of establishing correspondences between the elements of Petőfi’s communicative model, and on the other, the communicative process at the art museum, and gives a real example of a museum inspired by Semiotic Textology: Palazzo Buonaccorsi in the city of Macerata, where Petőfi was professor at the Faculty of Philosophy from 1989 until his retirement.

The scientific contributions are preceded and followed by two more personal texts. One is the biographical note written by Luciano Vitacolonna which offers an insightful depiction of how difficult it was to segregate academic from personal aspects in the relationship with Petőfi. The other is the lengthy interview with Petőfi by József Andor on his 70th birthday in 2002, which had been previously published in Hungarian, but has been translated into English and published for a much wider public for the first time here. This interview will give the reader the chance to ‘hear’ Petőfi’s voice talking about his life, his education, the first steps in his teaching and research career, his international experience, the birth of Text Linguistics, the research context at the University of Bielefeld—that became the main centre for textual research in Western Europe—, his last theoretical model, Semiotic Textology, and some of its key concepts

(cohesion, connexity, coherence, textuality), among other topics. This interview is an excellent complement to this book and, in our opinion, readers not yet acquainted with the figure and theory of Petőfi should begin the book by reading this interview.

In our presentation of the volume we have solely focused on how Petőfi's theories and ideas have been adopted, rethought, criticised or put into dialogue in different fields of research, yet the scientific contribution of these papers is much wider and richer, as the reader will soon discover. Our desire was simply to highlight the academic and personal picture of professor János S. Petőfi offered by the eleven texts gathered in this volume.

If we pay attention to the theoretical contributions mentioned by these authors alone, then at least we can say that the following topics have been touched on by one or several papers: the notion of coherence, the syntactic, semantic and pragmatic dimensions of his textual models, the model of multimedial communication process, the interdisciplinary network envisaged for his last textual model, the projection of his textual model in other disciplines such as Computational Linguistics and Rhetoric. This book, which is intended to be a tribute to the figure of János S. Petőfi, can only give a partial picture of the breadth and complexity of his theories. Nevertheless, the colleagues, disciples and friends of his that have taken part in it hope it will constitute an invitation to revisit his works and to revitalise his powerful thinking.

IN MEMORY OF JÁNOS S. PETŐFI

LUCIANO VITACOLONNA

Metaphorically speaking, I met János S. Petőfi for the first time in 1978, when I read two of his papers in *La linguistica testuale*, edited by M.-E. Conte. Literally speaking, however, it was on April 5, 1981 in Macerata that I first met the great Hungarian scholar. From then on, I had the honour and the pleasure to work together with him for more than thirty years.

Of course, when János moved from Bielefeld to Macerata in 1988, our meetings became more frequent and easier, our telephone communications longer and more regular, our shared work plans increased and, last but not least, our conversations became more intimate, so to say. Indeed, if up to the beginning of the nineties, we spoke mostly of text linguistics, semiotics, and the philosophy of language, after that time János became more and more willing to speak with me about his private life: his childhood; the untimely death of his father; the time spent working in a factory; the ill-famed year of 1956; his studies; and his fellow researchers with whom he was becoming more and more opposed and even, for ideological reasons, thwarted; his theatrical performances; his degrees and his work at the Computer Centre at the Academy of Science in Budapest; his leaving of his beloved Hungary and his stay in Sweden, and later Konstanz and Bielefeld; the beginning of the *Textlinguistik*; the origins of the *TeSWeST* and so on.

For privacy's sake, I can't dwell on these conversations; yet I want to stress the following: János was – so to speak – an oral encyclopaedist who could reconstruct, with lucid passion and absolute accuracy, all the main political, social and cultural events that had occurred over the course of many years. But listening to János also meant making the acquaintance of a wonderful storyteller: I was always struck by the intonations of his voice, by his pauses, by his Hungarian and German inflexions, by his narrative calm, by his firm self-confidence, and especially by his piercing glance aimed at recovering or fathoming the labyrinths of his experiences, of his *Erlebnisse*. I also remember the way János read the work of his favourite poets. But even though his memories were rich in pathos, the

narrative *Spannung* was often appeased by amusing anecdotes concerning linguists, semioticians and logicians, all told while eating *penne all'arrabbiata*.

If I can't dwell on János's private confidences, nevertheless I can say what János S. Petőfi has meant, and continues to mean, to me from both a scientific and a human point of view. From a scientific point of view, I consider János as a Master for many a reason: for his extraordinary working capacity as well as for his powers of concentration and reflection upon the foundations, organization and global aims of his theory as well as upon particular theoretical and methodological questions; for the formal exactitude, coherence and efficiency of this theory; for his ability to modify and adapt this theory to the requirements, expectations and aims brought about either by some incidental shortcoming noticed in the theory itself or by the unavoidable changes that occur in reality. Moreover, I consider János as a Master for his approach to that reality that lies beyond all schematism and dogmatism.

More specifically, I consider Semiotic Textology as the only theory that is able to tackle and explain the phenomenon of signification, communication and interpretation, since it is the only theory that can integrate the grammatical, semantic and pragmatic components within a semiotic framework, which involves different kinds of logic (formal logic, model-theoretic logic, etc.), different kinds of semantics (intensional and extensional semantics), different kinds of analyses (structural and procedural analyses), as well as different disciplines (linguistics, semiotics, philosophy, psychology, ethnology, and so on). Yet the basis of Semiotic Textology always has been and will always remain the text in its formal-semantic complexity and in its relationship with its communicative interlocutors. For this reason, great importance is to be attached both to the representation languages and to the atomic text.

From a personal and human point of view, I also consider János as a Master because he taught me to face life with seriousness but not with weightiness, with respect but not meekness, with irony and self-irony but not sarcasm or conceit. He was always a friend who was willing to listen and to help me, honestly and affectionately.

The last time I saw János was on October 13, 2012, in Budapest. At the time, I took some photos, but uncannily they all came out blurred. I realized then that it was a kind of foreboding: I would never again meet my beloved Master. Kevdes János, thank you.

LINGUISTICS

ON THE TEXT ORGANIZING STATUS OF KEY WORDS IN NEWS DISCOURSE: AN EMPIRICAL STUDY IN TERMS OF JÁNOS S. PETŐFI'S LEGACY

JÓZSEF ANDOR

ABSTRACT: This chapter provides an empirically-based analysis of the textual-functional role of key words in news discourse, relying on the now classic textological-interpretative model (text-structure/word-structure theory) created by János Sándor Petőfi, further extended in his multimodal, semiotic textological model. In the first part of the chapter, Petőfi's theoretical model is critically reviewed in comparison with parallel models of textual description and interpretation, mainly with those of Halliday and Hasan, de Beaugrande, van Dijk, and Widowson, with special emphasis on discussing the levelling of text and discourse in these models. Petőfi's three-partite theoretical model is given priority in the argumentation, in which criteria and phenomena of connexity (at the micro-level), cohesion (at the meso-level), and coherence (primarily at the macro-level of textual representation) were theoretically and empirically delineated. The second, main part of the chapter discusses the textual role of key words primarily functioning at the level of coherence and partially in cohesive lexical hierarchies, but also influencing and simultaneously relying on connexity-based textual representation. It is pointed out that key words are not statistically-, frequency-based textual units as hypothesized mainly by computational linguistically oriented scholars, but they have a text-organizing role and status, textually mapping mentally stored types of schematic information that are conceptually, hierarchically represented networks originating from frame-based, scenic-, and scriptal knowledge. In justification of hypothesizing the frame-semantic, conceptual basis of key words operating primarily at the level of textual coherence, an empirically-based, comparative-pragmatic analysis of the key words of two articles taken from same day issues of prominent media sources, The New York Times and The Independent, is presented based on their

selection by the software package WordSmith Tools and their identification by 30 adult native speakers of American English.

KEYWORDS: Text structure, coherence, lexicon, news discourse, keywords

1. Introduction

Systemic textology is a fairly young field of textual studies; its short history dates back only about 40-45 years. But even within the bounds of such a short period of time, models of textual description, influential or not, have been coming and going. The centrally important issue that almost all of them tackled concerned the mode and multiplicity of the levelling of representation of textual information coded from sentential strings into textual structures. Special attention has been devoted to identifying and studying the mode and nature of norms of textual representation at each level identified, described, and interpreted. To name just a few of the most influential models, we have (or have had) the Hallidayan-model in terms of the functionally bound, systemic linguistic framework of its founding father, Michael A. K. Halliday and his wife, Ruqaiya Hasan, pioneered by their 1976 book *Cohesion in English*, the structuralist semiotic approach of Algirdas J. Greimas (1987), the classic text linguistic work of Robert de Beaugrande and Wolfgang Dressler (1981), followed by several influential volumes and papers published by the first author of this popular introductory textbook (see, for instance, de Beaugrande 1984, 1997), the principally cognitivist, contextualist framework outlined by Teun van Dijk (one of the founding fathers of modern, systemic textology) in several of his volumes (see, for instance, 2008, 2014), some of them written in collaboration with his cognitive psychologist American colleague Walter Kintsch (Van Dijk and Kintsch 1983), Widowson (2004), Tanskanen (2006), all the way up to the influential work of our celebrated scholar, the Hungarian János Sándor Petőfi (especially 2004). As all of us celebrating him now very well know, he not only conducted pioneering research in text theory, developing his model, but he also coordinated the work of others around him, wherever he stayed, in Sweden, Bielefeld, Macerata, or Hungary, to name just a few of the major locations of his activities.

It is also a well-known fact that the history of the development of these models of textual description has been mainly linguistically oriented. Exceptions to such an orientation are comprised by only two of the founding giants of the field: those of van Dijk's socio-cognitively-based model and Petőfi's formally grounded, semiotic textology (let us not

forget about the fact that such an orientation on his part certainly owes a lot to his education in the field of mathematics). Another major innovation to the field on Petőfi's part was extension of the linguistically-based orientation to other types of textual representation, either by themselves or intertextually, thus leading to his comprehensive, complex model of multimodal textual semiotics. I believe that by this he markedly and significantly preceded the given type of research of his age. Let me just refer to brand new work on the textuality of coherence in a multimodal perspective recently published in a volume edited by Helmut Gruber and Gisela Redeker (2014), which, painfully, has not even a single reference to Petőfi's work.

As already mentioned, the issue of levelling textual representation and norms has been hotly debated since the early periods of modern, systemic textology. Most of the models of the early stages of the field delineated and circumscribed two levels of representation. In the classical model outlined by Halliday and Hasan (1976), local as well as global cohesion were identified and described. Under the scope of local cohesion, essentially, phenomena of grammatical nature (for instance, types of sentential linkage, spatial as well as temporal deixis, types of discourse markers) were identified and interpreted. Global coherence in the Hallidayan tradition mainly referred to the textual representation of semantic (both lexical and conceptual) types of cohesive relations. This latter notion, that of global cohesion, was then further extended and discussed under the term "coherence" by de Beaugrande, de Beaugrande and Dressler, van Dijk, van Dijk and Kintsch, and others. Unfortunately, though, the strictly linguistic relevance of the notion of coherence was retained in the work of most scholars, with the exception of van Dijk, and also that of János S. Petőfi. Another major innovation of Petőfi's was the introduction (and delineation) of a further level of textual representation, that of 'connexity'. In his model (2004), three levels, those of the micro-, meso-, and macro-level of textual representation were described. At the micro-level of linguistically coded texts, phenomena under the scope of "connexity" were identified, functionally representing subsentential, intrasentential, and suprasentential types of basically grammatical types of relation of the linkage of constituents. In his view, this is basically a morphematic level of textual representation (Petőfi and Benkes 1998: 58-59). At the meso-level of textual representation, semantically-grounded types of (primarily lexically-based) information are coded. It is at this level that the semantic matching of lexical collocation and priming, using the collocational potential of the lexical items textually represented, through a topical filter, is utilized by the text producer and decoded by the

comprehender. Petőfi, here, lays emphasis not only on the purely lexical semantic potential of words, but also on their encyclopaedic content, which latter facet of lexical representation significantly contributes to the pragmatic load and force of the lexis. (See his seminal study co-authored by Fritz Neubauer published as early as 1981 in Vol. 6. of his well-known textlinguistic book series *Research in Text Theory*.) Additionally, and perhaps unfortunately, we have to say, in its stage and status left unclarified, he also expressed views on the meso-level textual representation of syntagms (in his view a lexico-semantically, but also lexico-syntactically-based type of linguistic representation; cf. Petőfi and Benkes 1998: 58-59), paragraph structuring and blocking schemas, and the primarily lexical-semantic linkage between textual sentential strings (Petőfi 2004: 102). Due to his death five years ago, unfortunately, the clarification and analysis of the nature of cohesive linkage between the latter two, meso-level types of semantically and pragmatically based phenomena still lacks precise interpretation and analysis, leaving a lot of work to his students and colleagues. Even today, we do not and cannot have a clear understanding of why in his system he attributed the realization of the blocking schema and consequential paragraph structuring of linguistic texts both to the meso- and macro-levels of textual representation. All we can trace in his model (Petőfi and Benkes 1998: 128) is that the blocking schema and associated thematic propositional representation (Petőfi 2004: 51) are types of – what he called – macrocomposition, interfacing with the meso-level of representation of textual “sentential strings”.

In his model of the text-structure/world-structure theory, Petőfi attributes the highest status in the hierarchy of levels of textual representation and the decoding of textual content to “coherence”. In his view, the lower levels of textual representation and associated norms of adequacy satisfy conditions of intra-textuality, whereas coherence comprises an extra-textual, non-linguistically based level, wherein the mutual creativity of the decoder and interpreter gain ground. In his inaugural lecture given to the Hungarian Academy of Sciences in 2008, he stressed the view that even linguistic types of text cannot be described properly in purely linguistic terms (Petőfi [2008] 2014). Coherence is a creative, relational type of product of the communicating partners, wherein reliance on common ground (as outlined by Herbert Clark 1996), world knowledge in terms of conceptual, schematic structures like frames are grounded in discourse or the multi-modal representation of textual forms. Coherence is achieved by or in the inferential activity utilized in the process of communication (Petőfi 2008: 3). In achieving coherence, the

communicative partners, whoever and what they are, factually, rely on and utilize types of textual unificatory rules of the type identified by van Dijk under his term of “macrorules” (van Dijk 1980). In my electronic interview with János S. Petőfi in 2002 (see this volume for the English translation), he firmly expressed the view that connexity and cohesion do not serve as a necessary input to the achievement of coherence (Andor 2002: 581). I would like to add two further types of conceptual structure to the set of conceptual schematic structures: scenes and scripts. As explored in Andor (1985, 1989), scenes (and scenic knowledge) are global types of conceptual schemata with a horizontally structured makeup, and their accumulated content depends on directly gained experience, that is, scenes are grounded in direct exposure to world knowledge. On the other hand, frames are also globally and horizontally structured, but their content is based on indirectly gained types of world knowledge. The third type of conceptual structure, the notion of script (and that of scriptal knowledge), is global in nature, but it is vertically structured as a culturally dependent fixed set of schema, a sequence of typical actions in contextually dependent situations. As pointed out in Andor (1985) and also by various cognitive psychologists, the content of (and relations between) these three types of conceptual structure is not dependent on purely grounded categorization, but the types of schemata are usually merged onto one another in the course of cognitive development.

Although Petőfi had a clear mind on the operation of these conceptual structures concerning their role in contributing to the creation of coherence, he never managed to provide a thorough analysis of their role in text construction and/or interpretation in the semiotically based descriptions of his ample number of case studies of various types of multimodal textual representation. He probably did not live long enough to carry this out. On the other hand, he did point to their role concerning the textual status of thematic progression, the operation of which he attributed to the norms of macro-textual-level phenomena, that is, the domain of coherence. We do know, however, from his work (Petőfi 2004) and the research of others, that the facets, the formal realization of thematic progression are closely related to the norms of textual cohesion at least as much as they can be attributed to the level of coherence. Hence, an interface-related role of the body of norms of textual linkage operating at the meso- as well as the macro-level can be traced, pointing to the hierarchical character of the operation of textual norms which Petőfi revealed in his model. It can be clearly seen from the description of his model how the micro-level sets of norms, those of textual connexity, are strictly objectively grounded in linguistic representation, whereas, as we

get higher in the hierarchy of representation, to the level of coherence, its creation and realization (extra-textual, as has been noted above) are highly individualized, becoming a subjectivized matter. All this idea is reiterated in the Finnish scholar Tanskanen's (2006) research on the relation of textual cohesion and coherence. We can say that what we definitely miss from Petőfi's model is precise formulation and interpretation of the nature of textual perspective as well as the notion and nature of the process of perspectivization, for instance, the interpretation of perspectival shifts of / in discourse in his case studies.

Texts, for Petőfi (2004: 29), are relationally and semiotically grounded objects of informational representation. They cannot be described in purely linguistic terms, as they are not strictly and exclusively representations of the linguistic system. They are elements in the system of linguistic usage, deeply grounded in pragmatics.

2. The character, status and textual role of key words

Petőfi devoted considerable attention to investigating the textual role of words, the representation of the lexis in his theory. However, he mainly concentrated on the investigation of primarily strictly lexical-internal types of information: the relation of the lexical sense of the items and their encyclopedically grounded content from a textological point of view. He held the view that the thesaurus-based load, encyclopaedic information coded in the textual potential, force of words serves as an element of, a contribution to the realization of textual cohesion (Petőfi 1997: 54; 2004: 17). He failed to discuss the role of the lexis at the macro level, its coherence creating potential, that is, the force of lexical units to evoke and represent conceptually based, primarily frame-based knowledge. It is this issue which I would like to address briefly in this paper. Due to the linguistic representation of sense, words have the capacity to evoke and instigate frames and scenes from the conceptual lexicon. As such, their sense relations are mentally organized in sets of conceptual hierarchies. Items which have a central role in representing frames in the mind and which have sets of bodies of lexis under their control, conceptually related to them under the dominance of a frame, have a key status in the textual representation of their underlying frame. Such words are key words. Their potential for expressing keyness is hierarchically layered, that is, key words of a higher rank from the point of view of their frame-instigating potential may have lower ranked key words under their conceptually based control. Moreover, lexical items which potentially can serve as key words in representing particular frames may, in their textual representation, also

belong under the dominance of other frames in the same textual representation (for instance, a newspaper article on some topic). Words with such a potential lexical behaviour are called – using Herbert Clark’s terminology – bridging terms (Clark 1977). Along the grounds of this conceptually-based potential, key words certainly have a high degree of text-organizing potential. They are the items owing to which frames and scenes, as well as lexically coded scripts, are identified during textual comprehension and interpretation (Andor 1989, 2009). Such potential of the lexis exerted via encyclopaedic content (the importance of which for textual analyses was highly supported by Petőfi) is closely related to the potential of words called, in current research in lexical pragmatics, “lexical salience”. Largely differing from what is called “perceptual salience”, conceptual, cognitively-based, linguistically relevant salience critically enhances textual cohesion (via lexical priming, see Hoey 2005) as well as frame-instigated coherence (see Giora 2003 and Kecskés 2014: 176-199).

The view of keywordhood briefly outlined above (but discussed in detail in Andor 2009 and 2015) is radically different from the notion of keyness based on lexical statistics, the critical criterion of which is the degree of textual frequency of occurrence, the view advocated by computational and corpus linguist Mike Scott, founding father and developer of WordSmith Tools (Scott and Tribble, 2006). It is also radically different in its nature from the cultural embeddedness of the notion described in Raymond Williams’s pioneering work (1976), and Anna Wierzbicka’s and Cliff Goddard’s model of Semantic Metalanguage (the NSM theory), where key words are identified with reliance on their cultural-typologically valid frequency of occurrence and their notional status of potential use as semantic primitives (Wierzbicka 1997).

As early as 1976, the computational linguist Eugene Charniak noted the fact that textual representation instantiates conceptual frames, and that frames and the lexicon are complementary. It is noted in Petőfi’s theory that «elements of a text are decomposed into a sequence of more elementary units combined by rules of combination. To understand a text means to know the definitions of its constituting elements and their extensional interpretation» (Metzing 1981: 322). Additionally, it is noted by Metzing (1981: 322), with reference to Petőfi, that the frame may supply information missing in the text itself. Let me add to this that this is executed by the textual status of the frame-instigating key words of the text. All in all, as noted by Fetzer and Speyer (2014: 87), coherence means discourse connectedness within the realm of relevant frames and framing. They add that «discourse coherence does not lie in the discourse itself but in the minds of language users and is thus a socio-cognitive construct»