6.

Semiotic-textological approaches to literary discourse[1]

KÁROLY ISTVÁN BODA — JUDIT PORKOLÁB

In this study, we would like to select and review the main results of a few studies published in the previous volumes of the periodical *Officina Textologica* which, according to their specific research aims and objectives, carried out semiotic textological analysis of literary texts using or examining the approach of co-referential analysis. Our main concern is to overview the basic methods and formalism of co-reference analysis which have been gradually developed from the first volume of *Officina Textologica* and become a powerful tool to explore the textological structure and thematic composition of literary texts.

1. The theoretical background: János S. Petőfi's conception about the polyglot research program in textology / text linguistics realized in *Officina Textologica*

1.1

In the first volume of Officina Textologica, János S. Petőfi outlined a polyglot reseach program in textology and text linguistics which, since the first volume in 1997, has been thoroughly elaborated throughout the studies and essays of various researchers in the subsequent volumes of the periodical.

According to the author, "the linguistic-based research on texts has been gradually intensified since the 1960's. Because of the linguistic background, this particular field of research has been referred to as *text linguistics* on the one hand, and *text theory*, on the other hand (the latter emphasizing the theoretical nature of the research). In addition, we can find terms which contain 'discourse' instead of 'text' forming 'twin pairs' of the same term", i.e. to refer more or less to the same field of research (Petőfi 1997: 7). As regards the terminology used in the periodical *Officina Textologica*, both the term 'textology' and 'text linguistics' suggest the use of traditional language-based linguistic tools, but **the term 'textology' also involves the use of specific 'textological' tools, i.e. tools which can operate with the knowledge about the world**. It is especially important when we analyse the semantic aspects of texts. In this sense, the term 'text linguistics' can be considered as the linguistic component of textology.

In the first chapter entitled '*The disciplinary environment of text study. Text linguistics and textology in text research*', the author outlines the concept of 'textology' and 'text linguistics' in terms of their relationship with the other fields of research (see Fig 1).

In Fig 1, we can see the fields of research related to textology, the applications of textology, as well as the relationships of the various fields and applications illustrated here. "Because the textology of verbal texts in language L requires the text linguistics of verbal language L, which itself is built on the linguistics of the systems of verbal language L, and the linguistics of the use of the elements of those systems, the text linguistics of verbal language L provides a connecting link between the non-text-linguistic sector of linguistics (which, all things considered, is a sector of the grammar of system sentences), and the sector of textology" (ibid. 11). In Fig 1, the terms 'verbal texts' and 'verbal language L' refer to the fact that similar disciplinary environments can be (and should be) created for multimedia texts in the same way.

APPLICATIONS OF THE RESEARCH RESULTS OF TEXTOLOGY AND RELATED FIELDS, AND LINGUISTICS the domain of various types of texts: everyday texts, scientific, legal, religious, political, literary, etc. texts		
RELATED FIELDS TO TEXTOLOGY OF VERBAL TEXTS IN LANGUAGE "L" poetry, narrative, rhetoric, stylistics	TEXTOLOGY OF VERBAL TEXTS IN LANGUAGE "L"	LINGUISTICS OF VERBAL LANGUAGE "L" • linguistics of the (lexical and prosodic-rhythmic) system of verbal language "L", and linguistics of the use of the elements of its system • text linguistics of verbal language "L"
GENERAL POETRY, NARRATIVE, RHETORIC, STYLISTICS	GENERAL TEXTOLOGY	GENERAL SYSTEM AND TEXT LINGUISTICS OF VERBAL LANGUAGES
DISCIPLINES OF THE INTERDISCIPLINARY FOUNDATIONS OF TEXTOLOGY philosophy, psychology, sociology / anthropology, semiotics, communication theory, disciplines of formal and empirical methodologies		

Figure 1: The Disciplinary Environment of Text Study (abbrev.) (Petőfi 1997: 10)

It can be clearly seen from Fig 1 that a possible application of textology, and therefore *semiotic textology* as well, is the analysis of literary texts.

In the first volume of *Officina Textologica*, János S. Petőfi deals primarily with two major features of textuality: co-referential elements and co-reference relations, and the linear arrangement of text constituents.

1.2

In this study our main concern is the review of the possible applications of co-referential elements and co-reference relations in the analysis of literary discourse (the issue of the possible linear arrangements of text constituents is thoroughly discussed in another study of this volume). Therefore, it is worth outlining the main concepts and ideas about co-referentiality expounded in the first volume of *Officina Textologica* by János S. Petőfi. First let us see a brief definition of the terms: "It is well-known that **co-referential elements** are referred to as those language elements or phrases which might be different in their verbal manifestations but, according to the conviction of the reader or interpreter, they **all refer to the same entity or fact of the world the text describes**. According to this sense, co-reference relations are referred to as the relations between the supposed co-referential elements." (Petőfi 1997: 24)

In the first volume, several examples illustrate the co-referential elements and co-reference relations within selected texts. The first example is a passage from an essay by *Gyula Illés* about the famous Hungarian poet, *Endre Ady*. The essay is entitled '*About Ady's heritage*' (ibid. 25). In this passage, the chain of elements which have co-reference relation with the proper noun 'Ady' can be relatively easily revealed taking almost exclusively the lexical-syntactic relationships into account. In the example below, we illustrate the corresponding co-referential elements (in our case nouns, pronouns etc.) by underlining them. The text sentences are denoted by ordinal numbers in square brackets:

> [K00] About <u>Ady</u>'s heritage (extract)[2]
> [K01] <u>He</u> had an impact. [K02] <u>His</u> individual and exceptional ability to compose sentences, imagery, expressions, etc. has been imitated merely by mediocrity, but <u>he</u> was indeed the one <u>who</u> teached the next generation those things <u>he</u> had also inherited from <u>his</u> predecessors: the poet's courage and boldness to take on new thoughts and ideas, and to face the challenges. [...]

The second example is a passage from a short prose by *Lajos Áprily* entitled '*Birds above the sea*' (ibid. 26).

---

[2] (translated by BIK)

[K00] Birds above the sea (extract)[3]

[K01] On the bridge of the Danube, there are people standing near the bars, adults as well as children. [K02] They are looking at the gulls swinging above the river and, from time to time, landing on the water. [K03] In our reach of the Danube, too, I often see those birds with coral red beaks hovering over the water. [K04] Sometimes they fly high, and my eyes are loosing sight of them in the blueness or in the mist.

[K05] But I have seen gulls above the sea, too. [...]

In order to illustrate the co-reference relations in a more general sense, in this example János S. Petőfi takes into consideration the following four groups of co-referential elements:

0. conjunctions (including adverbial conjunctions of time)
1. elements which refer to the narrative self ("I")
2. elements which refer to the birds
3. elements which refer to the Danube and the sea.

Table 1 shows in its columns some co-referential elements of the passage belonging to the corresponding groups given above. As regards the question which elements belong to a particular group, it is important János S. Petőfi's remark (given in the fourth chapter) that the relationship of the elements within a group (e.g. birds, gulls, beaks etc.) "can only be explicitly established using a lexicon (i.e. a vocabulary having a specific structure) which also contains a thesauristic arrangement of its entries" (Petőfi 1997: 54)

text sentences	types of co-referential elements			
	0.	1.	2.	3.
[K00] Birds above the sea			**birds**	the **sea**
(end of paragraph / title)				
[K01] On the bridge of the Danube, there are people standing near the bars, adults as well as children.				(on the bridge of) the **Danube**

---

[3] (translated by BIK)

[K02] They are looking at the gulls swinging above the river and, from time to time, landing on the water.			at the **gulls**	(above) the **river** (on) the **water**
[K03] In our reach of the Danube, too, I often see those birds with coral red beaks hovering over the water.	too often	(in) **our** (reach of the Danube), **I** (often see)	**those birds** with coral red **beaks**	in our **reach** of the **Danube** (over) the **water**
(end of paragraph)				
[K04] Sometimes they fly high, and my eyes are loosing sight of them in the blueness or in the mist.	someti mes	**my** (eyes)	**they** [⇨ gulls]	
[K05] But I have seen gulls above the sea, too.	but too	**I** (have seen)	gulls	(above) the **sea**
(...)				

Table 1: Co-reference relations in a more general sense (slightly modified) (based on Petőfi 1997: 27-31)

As regards the importance of examining the various co-referential elements and their functions, "one of the tasks of text linguistics / textology is to analyse the functions that can be performed by words from different grammatical categories and by various syntactic structures as co-referential elements to ensure textuality." However, the author also emphasizes that "textuality can always be produced by several (linguistic and non-linguistic) factors that function together" (Petőfi 1997: 31).

The author lists certain lexical categories[4] that may serve a direct or indirect function to mark (potential) co-referentiality among text sentences. In Fig 2, the examples are taken from the extract from '*Birds above the sea*' given above.

---

[4] The categories of the parts of speech listed here are based from '*Brief Hungarian grammar*' by Endre Rácz and Etel Takács. (ibid. 32)

I. 1. The verb: stand, look at, see etc.

II. Nominal words:
1. The noun: the Danube, people, gull, sea etc.
2. The adjective: coral red etc.
3. The numeral: —
4. The pronoun: they, I etc.

III. Other (independent) parts of speech:
1. The adverb: there etc.
2. The infinitive: —; the participle: swinging, landing, hovering etc.
3. The interjection: —

IV. Not independent (auxiliary) parts of speech:
1. Verbal prefix: —
2. Postposition: —; preposition: above, over etc.
3. The article: the.
4. The conjunction: and, too, but etc.
5. The modifiers: —

Figure 2 Lexical categories functioning as (potential) co-referential elements

1.3

One of the most important aspects of the analysis of literary texts is to reveal their information content, that is, the information they convey. The basic problem is that "the great majority of texts contain only as much information expressed explicitly by the language, as is just enough for us to understand (i.e. to attach meaning to) the given text on the basis of our knowledge on the language and our knowledge about the world. If texts provided in every case all relationships, concerning the given language or part of the world that is referred to, by expressing them in a lexical way, they would become »unreadable«. But it is not necessary because we, the readers of the texts, are capable, on the basis of the syntactic/semantic information provided by the text itself, to think about the given text together with the relationships necessary to understand it in an almost automatic process — at least on the literal level of interpretation." (ibid. 34)

However, when we want to examine and describe co-referential elements and co-reference relations in a given text, it can certainly be useful — and sometimes necessary — to explicitly represent the information which can be obtained or deduced from the text on the basis of our linguistic or other knowledge. In his essay, János S. Petőfi demonstrates the nature of such information by examining a narrative passage from the Bible (ibid. 32).

The Birth of Moses (extract)
[1]Now a man of the house of Levi married a Levite woman, [2]and she became pregnant and gave birth to a son. When she saw that he was a fine child, she hid him for three months. [3]But when she could hide him no longer, she got a papyrus basket for him and coated it with tar and pitch. Then she placed the child in it and put it among the reeds along the bank of the Nile. [4]His sister stood at a distance to see what would happen to him. (Ex 2:1-4)
(The Bible. New International Version, 1984.[5])

When we try to reveal the information content of the text, the basic idea is to complete the text with words having individual (lexical) meaning in order to show the lexical references which are expressed implicitly by certain words (e.g. pronouns) or grammatical structures (e.g. the subject of sentences[6]) of the text. In addition, the completed text will be more compact and much easier to analyse if we introduce the so-called *co-reference indices* which replace, with corresponding codes, the persons, things, places etc. mentioned in the text. In the short passage given above, the list of co-reference indices can be as follows:

i01 = a man of the house of Levi,
i02 = a Levite woman married by i01 = the mother of i03,
i03 = the son of i01 and i02 = the ("fine") child = Moses
i04 = the sister of i03
...
i08 = the papyrus basket
i09 = the reeds
i10 = the Nile
i11 = the bank of i10
i12 = (at) a distance (from i11)
...

In addition to the introduction of co-reference indices, we shall use the concatenation (^) sign to mark the sequence ("concatenation") of words belonging to the same expression having an individual co-reference index. Moreover, we shall place in brackets those words and their indices which do not actually appear in the text but we insert them in order to reveal all the information that the text contains.

The Birth of Moses[i03] (completed with co-reference indices)
[1]Now   a^man^of^the^house^of^Levi[i01]   married

---

[5] http://niv.scripturetext.com/exodus/2.htm, 2012. 04.21
[6] Note that in the Hungarian language we should complete some sentences with either the subject or the direct object (or sometimes with both).

a^Levite^woman[i02], [2]and she[i02] became pregnant and (she[i02]) gave birth to a^son[i03]. When she[i02] saw that he[i03] was a^fine^child[i03], she[i02] hid him[i03] for three months. [3]But when she[i02] could hide him[i03] no longer, she[i02] got a^papyrus^basket[i08] for him[i03] and coated it[i08] with tar and pitch. Then she[i02] placed the^child[i03] in it[i08] and put it[i08] among the^reeds[i09] along the^bank[i11]^of^the^Nile[i10]. [4]His[i02]^sister[i04] stood at a^distance[i12]^(from^the^bank[i11]) to (she[i04]) see what would happen to him[i03]. (Ex 2:1-4)

The information we have added to the actual lexical manifestation of the given text as a form of co-reference indices ("replacements") can be deduced from purely grammatical knowledge. "These replacements can be performed because

- when we start reading a text, in our mind (technically speaking, in our mental model) the story becomes vivid: its entities (characters, objects, places etc.) appear and 'begin to live';
- in their first appearance, the entities are usually referred to by words or phrases having individual (lexical) meaning. Later, when they appear again, they might be referred to by other words, e.g. pronouns or determiners (or affixes in Hungarian, etc.); we are capable to establish the co-reference relationship between them and the words with individual meaning the entities have been first referred to by.

During the reading (or hearing) process we are always thinking in two dimension: first, in the dimension of text (looking for the relations between language elements without individual meaning and words with individual meaning); second, in the dimension of the relationship between the text and our mental model (looking for the relations between the referential elements of the text and the people, things, places, etc. in the mental model)." (ibid 33-34.)

Because our main concern here is about the analysis of literary discourse, we would like to make two important remarks which show the interpretation power of co-reference analysis.

First, in some cases grammatical (and/or linguistic) knowledge is not enough and we need knowledge about the world to complete the text with the necessary information. For example, on the basis of the very existence of Moses' sister (i04) who, in the Biblical passage, watched Moses when he was in the basket, we have to suppose that Moses was not the first child of his parents. Although it is obvious from our basic knowledge about the world, there is no hint in the passage that she (or any sister / brother of his) was born before Moses. But our knowledge is adequate — actually, it is well known from (the other parts of) the Bible that Moses had an older sister, Miriam, and an older brother, Aaron.

Second, a very important aspect of the selected passage from the Bible is that there are some elements *which cannot be understood from the text alone*. Our common knowledge (which is based on our everyday experiences) says that a sane woman will never get rid of her newborn child — unless she is compelled by extraordinary circumstances. This contradiction should be resolved by revealing what could have happened; it is the very point of the analysis when simple (i.e. intratextual) text linguistic methods are not sufficient any more and *intertextuality* is required. In fact, from the previous book of *Exodus* we can learn that the Egyptians were so afraid of the Israelites that "Pharaoh gave this order to all his people: »Every boy that is born [to the Hebrews] you must throw into the Nile, but let every girl live. «" (Ex 1:22)[7]

In the three examples presented above János S. Petőfi introduced three analytical approaches to illustrate co-referentiality in texts:

- the first example illustrated a single chain of co-referential elements underlined within the text;
- in the second example four different chain (or group) of co-referential elements were listed in separate columns of a table which itself represented the relationship between (the group of) co-referential elements and text sentences;
- in the third example all persons, objects and events which occurred (i.e. which were referred to) in the text were formally distinguished and represented by co-reference indices.

But, as the author says, "none of these examples can be considered an exact text linguistic / textological analysis because they lack the most important factor without which the analysis cannot be complete: to determine, on the basis of a thorough examination, whether using (text) linguistic methods alone is enough to explore co-referentiality or we should employ other — that is, textological! — methods as well" (ibid. 37-38).

2. The gradual development of the methods and formalism of co-reference analysis and its applications to the analysis of literary texts

2.1

In the second volume of *Officina Textologica*, giving a normative sample to the other authors of the volume, János S. Petőfi introduced and explained the basic formalism of co-reference analysis by selecting and analysing another passage from the Bible.

---

[7] The Bible. New International Version, 1984.
(http://niv.scripturetext.com/exodus/1.htm, 2012-04-25)

Ve:[8] The Calling of Matthew (Mt 9:9-13)

[9]As Jesus went on from there, he saw a man named Matthew sitting at the tax collector's booth. "Follow me," he told him, and Matthew got up and followed him. [10]While Jesus was having dinner at Matthew's house, many tax collectors and "sinners" came and ate with him and his disciples. [11]When the Pharisees saw this, they asked his disciples, "Why does your teacher eat with tax collectors and 'sinners'?" [12]On hearing this, Jesus said, "It is not the healthy who need a doctor, but the sick. [13]But go and learn what this means: 'I desire mercy, not sacrifice.'[9] For I have not come to call the righteous, but sinners."

(The Bible. New International Version, 1984.[10])

The main steps of co-reference analysis performed by János S. Petőfi can be summarized as follows (Petőfi 1998: 15-17):[11]

- the denotation of text sentences by **[Kxx]** (printed in bold type), where xx is the ordinal number of the corresponding text sentence;[12]
- the insertion of *additional information*, based on the text sentences themselves, their verbal context and/or our knowledge about the world, into the text sentences in order that they should be considered as individual information units. The inserted verbal phrases are presented in brackets and in italic (cursive) type;
  - when a verbal phrase is considered as one unit, its constituents are linked together by the concatenation (^) sign;
- the detailed explanation of the syntactic and co-referential analysis of text sentences is referred to as 'Commentaries' (not presented here but can be found in the complete co-reference analysis of the original study);

---

[8] 'Ve' is the abbreviation for 'Vehiculum' referring here to the selected passage to be analysed.

[9] „For I desire mercy, not sacrifice, and acknowledgment of God rather than burnt offerings." (Hosea 6:6)

[10] http://niv.scripturetext.com/matthew/9-9.htm, 2012-04-25.

[11] A more detailed description of the main steps of co-reference analysis can be found in another study of this volume by *Andrea Nagy* and *Franciska Skutta*.

[12] A text sentence is (by definition) begins with a capital letter and ends with a full stop (.), colon (:), exclamation mark (!) or question mark (?) (Petőfi 1998: 15). However, we would like to follow the steps of the original analysis of the selected passage presented by János S. Petőfi where the language of the passage was Hungarian. Because there are differences in punctuation in the Hungarian and English translations of the selected passage, in some cases we will not stick to the above definition when distinguishing text sentences.

- the list and description of *co-reference indices*; they are denoted by [ixx] where xx is the ordinal number of the corresponding index;
- the insertion of the co-reference indices into the text sentences,
  - placing an index immediately after the verbal phrase referred to by the index,
  - in case an index denotes a nominal phrase, its first occurrence is denoted by an equals sign (=) and in bold typeface.

The author remarks that "for the sake of simplicity, I will not deal with the analysis of articles, negative adjectives/adverbs, adverbial phrases of time, verb tenses and conjunctions" (Petőfi 1998: 16).

Now let us see first the co-reference analysis of the selected passage from the Bible, and then the list of the corresponding co-reference indices (cf. Petőfi 1998).

Ve/&ind:[13]

[K00] The Calling(i03)[i00] of Matthew[i03] (by Jesus[i01])

[K01] As Jesus[=i01] went^on[i01] from there[=i02], he[i01] saw[i01|i03] a^man^named^Matthew[=i03] (who[i03] was^)sitting[i03] at the^tax^collector's^booth[=i04].[14]

[K02] "(You[i03]) Follow[i03|i01] me[i01],"

[K03] he[i01] told[i01|"K02"] him[i03],

[K04] and Matthew[i03] got^up[i03] and (Matthew[i03]) followed[i03|i01] him[i01].

[K05] While Jesus[i01] was^having[i01|i09] dinner[=i09] at Matthew[i03]'s^house(i03)[=i05], many^tax^collectors[=i06] and (many^)"sinners"[=i07] came[i06,i07] (to the^house[i05]) and (they[i06,i07]) ate[i06,i07] with him[i01] and his[i01]^disciples(i01)[=i08] (in[15] the^house[i05]).

[K06] When the^Pharisees[=i10] saw[i10|"K05"[16]] this[K05], they[i10] asked[i10|"K07"] his[i01]^disciples(i01)[i08],

[K07] "Why does your[i08]^teacher(i08)[i01] eat[i01] with

---

[13] 'ind' is the abbreviation for 'indexed' referring here to the fact that the vehiculum has been analysed, indexed and completed by co-reference indices.

[14] Introducing a new co-reference index for 'tax collector' in general, a partial co-reference could be established with i06 (and with i03, etc.). Note that in the Hungarian translation the word 'custom' occurs instead of the 'tax collector's booth'.

[15] or (more probably) *in front of* the house (see later)

[16] János S. Petőfi's remark emphasizes that [K05] here "refers to 'a more general event' than 'this' in the given text sentence", but, for the sake of simplicity, we do not introduce another index to denote the corresponding part of [K05] which has been actually referred to (cf. Petőfi 1998: 23).

tax^collectors[i06] and 'sinners'[i07]?"[17]

[K08] On (Jesus[i01]) hearing[i01|"K07"] this[K07], Jesus[i01] said[i01|"K09^K10^K11^K12"] (to the^Pharisees[i10]),

[K09] "It is not the^healthy(^men)[=i11] who[i11] need[i11|i13] a^doctor[=i13], but the^sick(^men)[=i12] (need[i12|i13] a^doctor[i13]).

[K10] But (you[i10]) go[i10] and (you[i10]) learn[i10|i14] what[=i17] this[K11^K12] means["K11^K12"|i17]:

[K11] 'I[i01] desire[i01|i14] mercy[=i14], (I[i01] do^)not (desire[i01|i15]) sacrifice[=i15].'

[K12] For I[i01] have^not^come[i01] to (I[i01]) call[i01|i16] the^righteous(^men)[=i16], but (I[i01] have^come[i01] to (I[i01]) call[i01|i07] the^)sinners[i07]."

The list of co-reference indices that have been used is as follows:

i00the calling of Matthew[i03] by Jesus[i01] (did not occur in the original analysis)

i01Jesus

i02the place where Jesus[i01] was before the events of the passage

i03Matthew

i04the tax collector's booth

i05Matthew[i03]'s house

i06many tax collectors (that came to (i03)[i05])

i07many "sinners" (that came to (i03)[i05])

i08Jesus[i01]'s disciples

i09dinner (at (i03)[i05])

i10the Pharisees

i11the healthy (men), *in general*

i12the sick (men), *in general*

i13a doctor, *in general*

i14mercy, *in general*

i15sacrifice, *in general*

i16the righteous (men), *in general*

i17the meaning of [K11] and [K12] (not occurred in the original analysis)

As we can see, there are some additional rules which have been applied in the formal description of the text under analysis:

---

[17] Although in [K07] the phrase "taxt collectors and sinners" (and later, in [K12] the word "sinners") refers to tax collectors and sinners *in general* (i.e. to any of such persons), the author preferred the use of existing indices [i06] and [i07] (and [i07] in [K12], respectively). (cf. Petőfi 1998: 24)

(1) the co-reference index which denotes the *subject* of a particular verb should be placed right after the verb;

(2) the co-reference index which denotes the *direct object* of the verb is placed after the index denoting the subject, and the indices are separated by the vertical line (|) sign;

(3) when the subject (or the direct object) consists of more than one indexes (i.e. it can be expressed by a set of indexes), those co-reference indexes are separated by the comma (,) sign; when the subject or the direct object consist of more than one index of text sentences, those indices are separated by the concatenation (^) sign;

(4) the *possessive noun* should be placed before its possession in brackets.

It is very illuminating and thought-provoking to overview the occurrences of co-referential elements in text sentences illustrated in Table 2 (cf. ibid. 29-30). (For the sake of simplicity the verbs are displayed at those rows only which belong to the subject of the corresponding verb).

Although the author does not deal with the analysis of the text as literary discourse, it can quite possibly be stated that the co-reference analysis of the text reveals details that no other analytic method is likely to reveal. For example, it is worth observing that in the selected passage there are at least three levels of meaning:

(1) the description of the events which actually happened at the time of Matthew's calling;

(2) the communication between first Jesus and Matthew, then between the Pharisees and Jesus. Note that the indices of the quoted sentences, which are embedded into the analysed passage, occur in the same grammatical position where we otherwise use co-reference indices referring to "real" entities of the first level;

(3) the general (or abstract, symbolic, etc.) meaning of all (or some) of the events and related behaviour which have been transformed, by the questions of the Pharisees and the answers of Jesus, into the ethical rules of behaviour to be followed (including the 'right' or 'wrong' assessment of them).

Moreover, there is one point in the selected passage which, in case we perform only an intuitive analysis, is very easy to miss. It is the (quite unexpected) appearance of the Pharisees which does require an explanation. The necessary knowledge, when our basic knowledge about the world is not enough, might come from some *background knowledge,* e.g. from other texts. In our case, a significant text sentence from Luke can help: "Then Levi held a great banquet for Jesus at his house, and a large crowd of tax collectors and others

were eating with them." (Lk 5:29)[18] In the feast a great number of people participated, so it should have been outside the house of Levi (Matthew). Accordingly, the Pharisees, who were eager to examine every step of Jesus, could see the event, so they could make their comments.

---

[18] The Bible. New International Version, 1984. (http://niv.scripturetext.com/luke/5-29.htm, 2012-05-05)

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i01	Jesus went^on ——— he saw	me	he told	him	Jesus was^ having ——— him his^ disciples	his^ disciples	your^ teacher eat[i01]	(Jesus) hearing ——— Jesus said			I desire ——— I not (desire)	I have^ not- ^come ——— (I) call ——— (I) (have^ come) ——— (I) (call)
i02	there											
i03	a^man^ named^ Matthew ——— (who) (was^) sitting	(You) Follow	him	Matthew got^up ——— (Matthew) followed	Matthew's ^ house							
i04	the^tax^ collector's- ^booth											

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i05					Matthew's^ house _____ the^house _____ the^house							
i06					many^tax-^collectors came _____ (they) ate		tax^ collectors					
i07					(many^)"sinners" came _____ (they) ate		'sinners'					(the^) sinners
i08					his^ disciples	his^ disciples	your					
i09					dinner							
i10						the^ Pharisees saw _____ they asked		the^ Pharisees		(you) go _____ (you) learn		

	K01	K02	K03	K04	K05	K06	K07	K08	K09	K10	K11	K12
i11									the^ healthy(^men) need			
i12									the^sick- (^men) (need)			
i13									a^doctor (a^doctor)			
i14											mercy	
i15											sacrifice	
i16												the^ righteous- (^men)
i17										what		

2.2

In addition to János S. Petőfi's study, there are several studies in the second volume of *Officina Textologica* which selected a literary text to be analysed following János S. Petőfi's sample analysis. We would like to overview their contributions to the development of the formalism of co-reference analysis as well as the comments of *János S. Petőfi* and *Edit Dobi* (Petőfi-Dobi 1998) who provided a detailed discussion of all the studies published in the volume.

In their study, *Károly I. Boda* and *Judit Porkoláb* (Boda-Porkoláb 1998) selected a passage from the Revelation (Rev 21:9–23). Although the text is very rich in *intertextual references*, some of them are deliberately ignored because they do not have explicit co-reference relation to the other sections of the passage. However, the authors select one of such references ("the seven bowls full of the seven last plagues") and attach a co-referential index to it to demonstrate its importance. "The entity that belongs to the phrase 'the seven bowls full of the seven last plagues' does not occur in the selected passage any more; nevertheless, its selection as a co-reference index can be accounted for by the fact that it makes a reference to the antecedents of the selected passage (i.e. to the sequence of visions describing the final judgement in the Revelation)." (ibid. 37) The authors raise the problem that some of the co-reference relations could be lost if we indicate only the co-reference index of the subject and that of the direct object after the predicate (e.g. in the verb pattern sb_1 *tell* sb_2 *sth* the indirect object of the verb 'sb_2', although not indicated, might be important when referring to a person in the text) (ibid. 37). Another interesting aspect of the selected passage from the Bible is that it can be interpreted using different paradigms each having a different level of meaning which might (perhaps should) be reflected in the interpretation of the corresponding co-reference indices; e.g. New Jerusalem could be interpreted as the ancient Jewish city, the Church of Jesus, the human soul or God's Kingdom (ibid. 40).

In her study, *Márta Tuba* (Tuba 1998) selected a folk (*Palots*) tale for analysis. The author introduced several new denotations; they are as follows (Tuba 1998: 142-143):

- the possessive case is represented by connecting the possession and the possessor with the plus (+) sign (e.g. in the title of the tale, the^cock**[=i01]**'s^gold^coin[=**i02**+i01]);
- a change in the meaning of a co-reference index is indicated by the asterisk (*) sign (e.g. [i04] means the king as a person in the phrase *the king's brother*, compared with [i04*] which means his role or rank in the sentence *her brother was a king*);
- a co-reference is marked with an apostrophe (') when it is used to address somebody (e.g. i01 means 'the cock', i01' means the addressing

of the cock in the sentence 'My little cock, give me the gold coin!'). Note that "the referential value of addressing is a kind of 'language reference' which is different from the real world entity which might be referred to as the addressee" (Petőfi-Dobi 1998: 252).

It is worth noticing that "the most general organizing factor of the tale is the repetition" of various segments of the text (Tuba 1998: 161). For the exact identification of those segments the author uses the denotation K18a, K18b, etc. to mark the different clauses of the text sentence K18. To emphasize the importance of the indexing of time, the author remarks that "the most characteristic feature of the narrative type of texts is the sequence of events in time, which has been indicated by a unified index i00 to mark the different adverbial phrases of time" in the analysis; for example, 'once^upon^a^time[i00]', then[i00], in^three^days[i00], etc. (ibid. 161). This index does really express the dimension of time but "does not represent the progress of time" (Petőfi-Dobi 1998:259).

In his study, *Imre Békési* emphasizes that "if the sufficient number of analyses have been performed and their descriptive apparatus is unified to the extent that their results can be coded for computer processing, then we can seriously hope for new discoveries" (Békési 1998: 185). His starting point is the so-called 'thesis sentence' which, in some type of texts, comprises (either before or after, and in rare cases in the middle of, the corresponding sentences) the essence of the content that the text describes. In the selected passage from *Anna Karenina* by Leo Tolstoy, the first sentence can be considered as a thesis sentence which contains all the co-reference indices which occur in the following text sentences of the passage: "The mother cannot even compare Vronsky with Levin (*because she thinks the characters of the two men so different*)." Note that the completion of the text sentence with explanatory information (given in brackets) seems to be very important because the organization of text sentences "is realized under the control of contextual-logical (semantic) relations as well as communication, rhetoric etc. rules" (ibid. 191). As a consequence, the co-reference relations of the text sentences of the selected passage can be organized (or illustrated) by the use of conjunctions such as 'because', 'therefore', and 'however' (ibid. 192-193).

In her study, *Katalin B. Fejes* analyses the poem '*Greeting Thomas Mann*' by the famous Hungarian poet, *Attila József*. Her approach to the poem is determined by the fact that the poem "is actually a greeting, therefore its creation could have been affected by the conceived situation where the poem was to be read out loud. This situation could have been so effective that its influence was imprinted on the poem." (B. Fejes 1998: 194) According to the main senses, the author distinguishes three different communication channels each containing a corresponding chain (or "bundle") of co-referenced lexemes from the poem:

auditive channel	kinetic channel	visual channel
[K01] tell a tale (whether he wishes) the tale tell a tale [K02] Tell us (what ...) tell a tale (about ...) [K03] tell us the truth [K05] your gentle voice tell a tale (about ...) [K06] you can read out loud [K07] if you speak [K09] let's start telling a tale [K10] We are listening to you	[K01] Please, don't go (or he wishes) you to be there Sit down among us [K02] you are here with us [K08] Please take a seat	[K10] who is just looking at you (who) sees a European

Table 2: Three communication channels and the co-referenced lexemes belonging to them (ibid. 195; slightly modified)

The author assigns co-reference indices to (some of) the nominal and (some of) the verbal elements of the selected text, as well as "to the referential values of the statements" (Petőfi-Dobi 1998: 260), focussing on the auditive-kinetic-visual relationships represented by the given channels. These relationships play a central role in the interpretation of the poem because "the joy of greeting can hardly be expressed by mere words, so the 'we' are expecting words from the 'you': ask, request, question, or look, rejoice, see. Therefore we might suppose

that the composition of the whole text is centered around the auditive-kinetic-visual representation of the very moment of 'we ask you'." (ibid. 205)

Interestingly enough, the composition of the poem also contains some outside references. There are some elements in the text which, although do not have explicit co-reference relationship with the other elements, refer to certain knowledge outside the textual world of the analysed text (e.g. the two main characters of *The Magic Mountain* by *Thomas Mann*, the funeral of the famous Hungarian poet, *Dezső Kosztolányi*, etc.) (ibid. 205). Although the author does not explicitly mention, to represent those references we might as well add another channel concerning memories and imagination; i.e. some "mental content" that can be associated with the situation the poem describes. This shows that the idea of identifying co-reference bundles (i.e. co-reference chains the elements of which refer to a central theme or topic of the text) can be a powerful tool for the interpretation of a literary text, fully supported by the formalism of co-reference analysis.

As regards the formalism used by the author, in the co-reference structure of the analysed text we can observe "full" as well as partial or thesauristic co-reference relationships between the corresponding lexemes within each bundle (e.g. the first occurrence of '*tell a tale*' in K01, where a child asks its parent(s), has a partial reference to '*tell a tale*' in K01 / K02 etc. where the 'we' ask the 'you' in the textual world of the poem; note that the subsequent occurrences of '*tell a tale*' are in "full" co-reference relation with each other referring to the same general entity and/or action; an example for the thesauristic co-reference relationship might be 'you *speak*' vs. 'we are *listening* to you'). Note that the issue of the identification and/or denotation of the partial co-reference relationship also arises in other studies (e.g. in Tolcsvai 1998: 229, Petőfi-Dobi 1998: 261, etc.).

In his study, *Gábor Tolcsvai Nagy* analyses a selected part of a so-called 'dictionary' novel "*Dust*" by *Ferenc Temesi*. The selected text has a form of a dictionary entry having the headword '*ideal*' (or '*Dust.ideal*' to indicate the place of the 'headword' within the 'dictionary'). It includes several references which establish connection with the other parts of the dictionary novel and therefore can only be fully understood with the comprehensive knowledge of the full text. The co-reference analysis made by the author, following the basic methodology and formalism of János S. Petőfi's normative sample (Petőfi 1998), is intended to be an element of "developing a textological representation language" (Tolcsvai 1998: 223).

The analysis has several aspects that are worth considering. As an example, let us see the co-reference analysis of the fourth text sentence where the

narrative self of the selected text, who in his youth[19], 'literally' chose Percy Harrison Fawcett as his ideal, expresses his deep emotions:

> [K04/&ind] In it,[20] this valiant^British^traveller[i06] related his[i06] adventures in South America and he[i06] suffered so many sore trials in the jungle of Amazonas that, considering Zola or not, it [he? they?; see the 'commentary' below] really touched[i08|i03] me deeply.[21]

i03the narrative self, "I"
i06Percy Harrison Fawcett, the "valiant British traveler"
i08the destiny or behaviour of Percy Harrison Fawcett

Note that several co-reference indices are deliberately missing. As the author says, "the analysis is not 'complete' because it focuses, in the first place, on the most uncommon nominal parts, verbal arguments and pronominal elements which need to be explicated and have co-reference relation" with other elements in the selected text (Tolcsvai 1998: 225). One of the most interesting aspects of the analysis of the text sentence K04 is that the subject of the last clause ('it') is "not entirely unambiguous", partly because the original Hungarian text allows other translations as well (e.g. 'he' or 'they'). We might assume, however, that the entity that the subject refers to is, most likely, the destiny or behaviour of the traveller. "According to this interpretation, the subject of 'touched' is in partial co-reference relation with the 'adventures' and 'trials' constituents of the text sentence K04" (ibid. 229).[22] This calls our attention to the significance of the phenomenon of partial co-reference relationship which has also been mentioned in the overall discussion of the volume (cf. Petőfi-Dobi 1998: 261).

2.3

The *Epilogue* of the 2nd volume of *Officina Textologica* is an overall discussion and analysis of the main concepts, problems and further questions of co-reference analysis. Reflecting on some issues raised in the studies of the same

---

[19] as we learn from the first text sentence of the analysed text,

[20] That is, in the book first mentioned in the second text sentence of the analysed text („I got his book in Christmas"). Although it has a separate co-reference index (i07) but the author does not use it in the co-reference analysis of the fourth text sentence.

[21] (translated by BIK)

[22] Note that we can reasonably assume that one of the most important feature of an 'ideal' is his or her *behaviour* in certain, most likely difficult, situation. In this case, i08 (=the destiny or *behaviour* of Percy Harrison Fawcett) would establish a partial co-reference relation with the title or 'headword' of the 'dictionary entry' text ('ideal') which, by definition, should be the most significant word (a keyword or a kind of 'thesis word', cf. Békési 1998) of the selected text.

volume, *János S. Petőfi* and *Edit Dobi* overview and summarize the basic thoughts and questions of co-reference analysis as follows:

Co-reference analysis is performed by considering the relationship of two different levels (see Petőfi-Dobi 1998: 238):

- the first level is "the mental image of a particular segment of the world which has been expressed, according to the interpreter, in the text analysed"; as a consequence, this level contains the entities of the given world segment (i.e. persons, objects, concepts, attributes, statements etc.) and their corresponding mental images;
- the second level is "the vehiculum (i.e. the physical appearance) of the text analysed including the constituents of the text to be analysed semantically / syntactically by the interpreter. Some of these constituents refer to the entities of the represented world segment, while the other constituents express certain statements on the referred entities.

As regards the interpretation of co-reference relation, it is "a kind of relationship interpreted in some way between text constituents referring to the (supposedly) same entity of the represented world segment" (ibid.). Various types of this relationship can be distinguished, e.g. we can speak of 'identity' relationship in case all constituents refer to the same entity, or of 'possessor-possession' relationship when one referred entity (e.g. an adventure, destiny, behaviour etc.) belongs to the other referred entity (e.g. a traveller). Naturally, "there can be other types of relationship, too" (ibid.).

The authors list four very important questions to be answered when discussing 'co-reference analysis' in general:

(1) "what kind of entities can be referred to by the constituents of text";

(2) "what kind of text constituents can have reference" to the above entities;

(3) "what types of co-referentiality can be distinguished, that is, what kind of referential difference" between the referred entities can be accepted so that their relationship can still be considered co-referential;

(4) "what can be considered as the most effective way to denote co-reference relations" (ibid. 238).

As to the detailed discussion of and possible answers to these questions we refer to the chapter of this volume written by Andrea Nagy and Franciska Skutta.

As the review and discussion of the main results of the above studies presented here most probably show, the formalism and methodology of co-reference analysis can be a very effective and highly analytic starting point to the semiotic-textological analysis of texts which in turn makes a substantial and valuable contribution to the analysis of literary discourse as well.

3. Further readings: a selected bibliography

In addition to the studies reviewed before, we have compiled a short bibliography of studies analysing literary texts using (or referring to) the concepts and methods of co-referential analysis in the subsequent volumes of Officina Textologica. They are as follows:

Bencze Lóránt: A koreferáló elemek és viszonyok retorikai-stilisztikai megvilágítása. [Rhetorical-stylistic approach to coreferential elements and co-reference relations.] Officina Textologica 4/2000, 35–56.

Benkes Zsuzsa: Kreatív gyakorlatok a koreferenciarelációk analitikus elemzésének előkészítéséhez. [Creative exercises as preparation for the analytical examination of co-reference relations.] (Példaszöveg: Biblia. Mózes II. könyve. 2. rész 1-11.) Officina Textologica 4/2000, 57–80.

Boda I. Károly, Bodáné Porkoláb Judit: A koreferencia kérdései a számítógépes szövegfeldolgozás szempontjából. [Problems of co-reference in computer-based text processing.] Officina Textologica 4/2000, 150–180.

Boda I. Károly, Porkoláb Judit: Téma-réma kapcsolatok vizsgálata egy kiválasztott versszövegben korreferenciaelemzés segítségével. [Examination of theme-rheme relations in selected texts using co-reference analysis.] Officina Textologica 7/2002, 93–112.

Boda I. Károly, Bodáné Porkoláb Judit: A tematikus progresszió vizsgálata kiválasztott szövegekben korreferencia-elemzés segítségével. [Examination of thematic progression in selected texts using co-reference analysis.] Officina Textologica 9/2003, 45–58.

Boda I. Károly, Bodáné Porkoláb Judit: Egy angol vers és magyar fordításainak összevetése a korreferencialitás szempontjából. [Comparing the co-referentiality of a poem by T.S. Eliot and its two Hungarian translations.] Officina Textologica 12/2005, 83–100.

Boda I. Károly, Bodáné Porkoláb Judit: Korreferencialitás és hipertextuális kapcsolatok a poétikai kommunikációban. [Co-referentiality and hypertextual relations in poetic communication.] Officina Textologica 16/2011, 117–129.

Csűry Andrea: Korreferenciaviszonyok vizsgálata magyar és francia dialogális szövegben. [Examination of co-reference relationships in Hungarian and French dialogues.] Officina Textologica 12/2005, 19–34.

Dobi Edit: A koreferenciaelemzés kérdésének rendszernyelvészeti megközelítése szemiotikai textológiai keretben. [A system linguistic approach to co-reference analysis in semiotic textological framework.] Officina Textologica 4/2000, 81–123.

Dobi Edit: A korreferencialitás megnyilvánulásai magyar és német szövegekben. [The manifestations of co-referentiality in Hungarian and German texts.] Officina Textologica 12/2005, 101–116.

Tolcsvai-Nagy Gábor: Kérdések a koreferenciáról. [Questions about co-reference.] Officina Textologica 4/2000, 11–34.

Table 4 shows all the studies mentioned above as well as the corresponding literary text (or a certain passage of it) selected for analysis by the authors with a view to using the methods and formalism of co-reference analysis, and/or examining its main issues and problems in general.

Study from *Officina Textologica*	**Selected literary text and genre**
B. Fejes Katalin: Kompozícióalkotó korreferencia-nyalábok. [Composition-forming co-reference bundles.]	Attila József: Thomas Mann üdvözlése [Greeting Thomas Mann] (poem)
Békési Imre: A korreferenciarelációk és a tételmondat. [Co-reference relations and the thesis sentence.]	Lev Tolsztoj [Leo Tolstoy]: Anna Karenina (novel)
Bencze Lóránt: A koreferáló elemek és viszonyok retorikai-stilisztikai megvilágítása. [Rhetorical-stylistic approach to coreferential elements and co-reference relations.]	Sándor Petőfi: Akasszátok fel a királyokat! [Hang the Kings!] (poem)
Benkes Zsuzsa: Kreatív gyakorlatok a koreferenciarelációk analitikus elemzésének előkészítéséhez. [Creative exercises as preparation for the analytical examination of co-reference relations.]	The Bible. Old Testament. (Ex 2:1-10)
Boda I. Károly, B. Porkoláb Judit: Koreferenciális kifejezések és koreferenciarelációk. [Co-referential phrases and co-reference relations.]	The Bible. New Testament. (Rev 21:9-23)
Boda I. Károly, Bodáné Porkoláb Judit: A koreferencia kérdései a számítógépes szövegfeldolgozás szempontjából. [Problems of co-reference in computer-based text processing]	The Bible. New Testament. (Rev 1:1-5)
Boda I. Károly, Porkoláb Judit: Téma-réma kapcsolatok vizsgálata egy kiválasztott versszövegben korreferenciaelemzés segítségével. [Examination of theme-rheme relations in selected texts using co-reference analysis.]	Milán Füst: A szőlőműves [The Vine-dresser] (poem)

Boda I. Károly, Bodáné Porkoláb Judit: A tematikus progresszió vizsgálata kiválasztott szövegekben korreferencia-elemzés segítségével. [Examination of thematic progression in selected texts using co-reference analysis.]	Miklós Radnóti: Lomb alatt [Under the Leafy Boughs] (poem)
Boda I. Károly, Bodáné Porkoláb Judit: Egy angol vers és magyar fordításainak összevetése a korreferencialitás szempontjából. [Comparing the co-referentiality of a poem by T.S. Eliot and its two Hungarian translations.]	T.S. Eliot: Marina (poem)
Boda I. Károly, Bodáné Porkoláb Judit: Korreferencialitás és hipertextuális kapcsolatok a poétikai kommunikációban. [Co-referentiality and hypertextual relations in poetic communication.]	T.S. Eliot: The Waste Land (poem)
Csűry Andrea: Korreferenciaviszonyok vizsgálata magyar és francia dialogális szövegben. [Examination of co-reference relationships in Hungarian and French dialogues.]	János Háy: A Gézagyerek [The Géza-boy] (play)
Dobi Edit: A koreferenciaelemzés kérdésének rendszernyelvészeti megközelítése szemiotikai textológiai keretben. [A system linguistic approach to co-reference analysis in semiotic textological framework.]	Lajos Áprily: A hiúz [The Lynx] (short story)
Dobi Edit: A korreferencialitás megnyilvánulásai magyar és német szövegekben. [The manifestations of co-referentiality in Hungarian and German texts.]	Eduard Mörike: Mozart auf der Reise nach Prag [Mozart's Journey to Prague]
Petőfi S. János (1997): Egy poliglott szövegtani-szövegnyelvészeti kutatóprogram. [A polyglot textological / text linguistic research program.]	Illyés Gyula: Ady örökségéről [About Ady's heritage] (essay) Lajos Áprily: Madarak a tenger felett [Birds Above the Sea] (short prose) The Bible. Old Testament. (Ex 2:1-10)
Petőfi S. János: Koreferenciális elemek és koreferenciarelációk. [Co-referential elements and co-reference relations.]	The Bible. New Testament. (Mt 9:9-13)

Petőfi S. János, Dobi Edit: Utószó. [Epilogue.]	(all the texts which have been selected by the studies in the 2nd volume of Officina Textologica)
Tolcsvai Nagy Gábor: Explicitség és koreferencia. [Expliciteness and co-reference.]	Ferenc Temesi: Por [Dust] (novel)
Tolcsvai-Nagy Gábor: Kérdések a koreferenciáról. [Questions about co-reference.]	Ferenc Temesi: Por [Dust] (novel)
Tuba Márta: Koreferenciális kifejezések és koreferenciarelációk egy mesében. [Co-referential phrases and co-reference relations in a folk tale.]	A kakas aranygarasa [The cock's gold coin] (folk tale)

Table 3: Selected studies from Officina Textologica and the literary texts analysed by them

References

B. Fejes Katalin (1998): Kompozícióalkotó korreferencia-nyalábok. [Composition-forming co-reference bundles.] Példaszöveg: József Attila: Thomas Mann üdvözlése. Officina Textologica 1998/2, 194–207.

Békési Imre (1998): A korreferenciarelációk és a tételmondat. [Co-reference relations and the thesis sentence.] Példaszöveg: Lev Tolsztoj: Anna Karenyina. Első rész, XII. 3-4. Officina Textologica 1998/2, 185–193.

Boda I. Károly, B. Porkoláb Judit (1998): Koreferenciális kifejezések és koreferenciarelációk. [Co-referential phrases and co-reference relations.] Példaszöveg: Szent János Apostol Jelenéseinek könyve. 21:9-23. Az új Jeruzsálem. (Részlet) Officina Textologica 1998/2, 32–56.

Petőfi S. János (1997): Egy poliglott szövegtani-szövegnyelvészeti kutatóprogram. [A polyglot textological / text linguistic research program.] Officina Textologica 1997/1. Debrecen: Kossuth Egyetemi Kiadó.

Petőfi S. János (1998): Koreferenciális elemek és koreferenciarelációk. [Co-referential elements and co-reference relations.] Példaszöveg: Mt. 9,9-13. Máté meghívása. Officina Textologica 1998/2, 15–31.

Petőfi S. János, Dobi Edit (1998): Utószó. [Epilogue.] Officina Textologica 1998/2, 238–261.

Tolcsvai Nagy Gábor (1998): Explicitség és koreferencia. [Expliciteness and co-reference.] Példaszöveg: egy anekdotikus prózarészlet. Officina Textologica 1998/2, 223–237.

Tuba Márta (1998): Koreferenciális kifejezések és koreferenciarelációk egy mesében. [Co-referential phrases and co-reference relations in a folk tale.] Példaszöveg: a kakas aranygarasa. (Részlet) Officina Textologica 1998/2, 142–168.