REPRESENTATION AND ANALYSIS OF
A CORPUS OF LEIBNIZ-CLARKE LETTERS

Ştefan TRAUŞAN-MATU[1], Tudor FOCA[2]

Abstract. *The paper presents an XML annotation structure for corpora of letters, and a series of processings for complex visualizations, obtained with XSLT transformations and with the RederBench environment. The underlying idea is the usage of the polyphonic model of discourse, which is very well suited for capturing the weaving of threads of discussions involving debates. As a case study was taken the correspondence between Leibniz and Clarke.*

Keywords: annotated corpora, natural language processing, polyhponic model, discourse analysis, Republic of letters

1. Introduction

In Europe of the 16-18 centuries, many philosophical and scientific debates have taken place by exchanging letters in what was called the "Republic of letters" ("Respublica literaria").

In recent years, the interest in analyzing these debates has increased. For example, at least two major projects have this subject, at Stanford University in the US (http://republicofletters.stanford.edu/) and in a COST action in the EU (http: www.republicofletters.net/).

A powerful tool for the investigations on the Republic of letters is information technology and, particularly, Natural Language Processing (NLP). In this direction, this paper presents a framework for the analysis of the correspondence-debate of ideas, and its implementation using NLP, based on the polyphonic model [1], inspired from Mikhail Bakhtin's ideas [2].

As a case study, the correspondence between Leibniz and Clarke [3] is used. These debates of ideas are very important and interesting because they implicitly involve Isaac Newton, whose disciple is Clarke, the discussions therefore being directly related to Newton's ideas.

Consequently, the correspondence reflects the confrontation of the ideas of two titans of science and philosophy, including topics such as God, space, time, soul, miracles, nature, etc. [3].

---

[1]Member of A.O.S.R. Prof., PhD, Faculty of Automatic Control and Computers, "Politehnica" University of Bucharest and Senior Researcher, Research Institute for Artificial Intelligence of the Romanian Academy, Romania. stefan.trausan@cs.pub.ro.
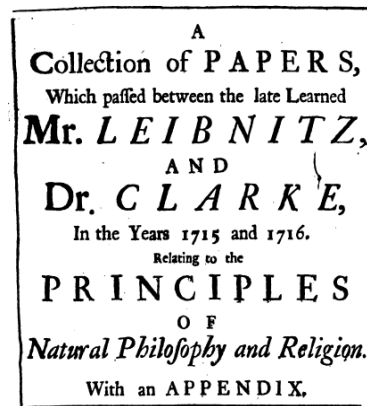[2]Student, Colegiul Naţional „Gheorghe Lazăr".

A
Collection of P A P E R S,
Which paſſed between the late Learned
Mr. *L E I B N I T Z,*
A N D
Dr. *C L A R K E,*
In the Years 1715 and 1716.
Relating to the
P R I N C I P L E S
O F
Natural Philoſophy and Religion.
With an A P P E N D I X.

Fig. 1. A volume containing the exchanged letters between Leibniz and Clarke [4].

2. The corpus

The corpus of correspondence contains eight letters, four authored by Leibniz, who starts the correspondence, and four written by Clarke. After the first letter, each of them has direct references to the previous one. Even if he did not authored any of the letters, Isaac Newton is a participant at the debate implicitly, through the ideas attributed to him and advocated by Clarke.

The corpus has been annotated in two phases. The text of the corpus was firstly semi-manually annotated according to the Document Type Definition (DTD) used for conversations [5], which included only annotations for turns (in our case, letters) and utterances (according to the paragraphs of the text). The references made in text were coded as the explicit references in the case of chats.

As source for the text was considered the version at
          http://www.earlymoderntexts.com/assets/pdfs/leibniz1715_1.pdf,
last accessed at 22 October 2017.

In a second phase, new annotations were introduced in order to code the main subjects of the discussion and the indirect utterances ("iutt"), that means utterances that are emitted by a person, but attributed to another. This latter element has an attribute for polarity (positive or negative).

An example of using these new elements of annotation is the following, where Leibniz writes what Newton said:

```
<Turn nickname="Leibniz" nr="1">Leibniz's first paper (November 1715)

<Utterance genid="1" ref="-1"> 1 Natural religion seems to be greatly on the
decline  -in  England-,  where  many  people  hold  that  human
<subject>soul</subject>s  are  made  of  <subject>matter</subject>,  and  others
contend that <subject>God</subject> himself is a corporeal being, -i.e. a body-.
</Utterance>
```

<Utterance genid="3" ref="-1">3 <iutt pol="positive">Newton says that <subject>space</subject> is an organ-like a sense-organ- by which <subject>God</subject> senses things. <c w="disj">But</c> if <subject>God</subject> needs an organ to sense things by, it follows that they don't depend entirely on him and weren't produced by him.</iutt>

```
▼<Dialog team="1">
  ▼<Participants>
    <Person nickname="Leibniz"/>
    <Person nickname="Clarke"/>
  </Participants>
  Leibniz-Clarke papers G. W. Leibniz and Samuel Clarke Clarke 1: (26.xi.1715)
  ▼<Body>
    ▼<Turn nickname="Leibniz" nr="1">
      Leibniz's first paper (November 1715)
      ▼<Utterance genid="1" ref="-1">
        1 Natural religion seems to be greatly on the decline -in England-, where many people hold that human
        <subject>soul</subject>
        s are made of
        <subject>matter</subject>
        , and others contend that
        <subject>God</subject>
        himself is a corporeal being, -i.e. a body-.
      </Utterance>
      ▼<Utterance genid="2" ref="-1">
        2 Locke and his followers aren't sure whether the
        <subject>soul</subject>
        is material and naturally perishable.
      </Utterance>
      ▼<Utterance genid="3" ref="-1">
        3
        ▼<iutt pol="pos">
          Newton says that
          <subject>space</subject>
          is an organ-like a sense-organ- by which
          <subject>God</subject>
          senses things.
          <c w="disj">But</c>
          if
          <subject>God</subject>
          needs an organ to sense things by, it follows that they don't depend entirely on him and weren't produced by him.
        </iutt>
        ▼<!--
          Clarke  translates Leibniz as speaking of how <subject>God</subject> &apos;perceives&apos; things; <c w="disj">but</c>
          -->
      </Utterance>
      ▼<Utterance genid="4" ref="-1">
        4 Newton and his followers also have a very odd opinion regarding
        <subject>God</subject>
        's workmanship. According to them,
        <subject>God</subject>
        's watch -the universe- would stop working if he didn't re-wind it from
        <subject>time</subject>
        to
```

Fig. 2. An example of annotation.

The full DTD used in annotation is:

```
<!ELEMENT Dialog ( #PCDATA | Participants | Turn )* >
<!ATTLIST Dialog team NMTOKEN #REQUIRED >

<!ELEMENT Participants ( Person+ ) >
<!ELEMENT Person EMPTY >
<!ATTLIST Person nickname NMTOKEN #REQUIRED >

<!ELEMENT Turn ( #PCDATA | Utterance )* >
<!ATTLIST Turn author NMTOKEN #IMPLIED >
<!ATTLIST Turn nickname NMTOKEN #IMPLIED >
<!ATTLIST Turn nr NMTOKEN #REQUIRED >

<!ELEMENT Utterance ( #PCDATA | c | iutt | page | ref | subject )* >
<!ATTLIST Utterance genid NMTOKEN #REQUIRED >
<!ATTLIST Utterance ref NMTOKEN #REQUIRED >
```

```
<!ELEMENT c ( #PCDATA ) >
<!ATTLIST c w NMTOKEN #FIXED "disj" >

<!ELEMENT iutt ( #PCDATA | c | ref | subject )* >
<!ATTLIST iutt pol ( neg | pos ) #REQUIRED >

<!ELEMENT subject ( #PCDATA ) >
<!ATTLIST subject attr CDATA #IMPLIED >
```

An example of the full annotation is presented in the Figure 2.

3. Analysis of the corpus

3.1 The polyphonic model

The correspondence between members of the Republic of letters includes debates on complex subjects related to philosophy, science, and religion as in the case of the letters exchanged by Leibniz and Clarke. Moreover, discussions usually include many times ideas of other people (for example, Newton in the previously mentioned case) or influential mentalities. Therefore, a deep analysis of the discourse is needed and, as, in our opinion, classical discourse analysis in NLP is not offering suitable tools; we considered that the advanced polyphonic model of discourse [1, 6, 7] is needed.

The polyphonic model is based on the theories of the Russian philosopher Mikhail Bakhtin [2]. He considers that dialog is omnipresent in our lives, that multiple voices are present in any text, even in a word, that the ventriloquism phenomenon occurs (one person speaks with the voice of another), and that sometimes voices weave in a polyphonic way. In our extension of the ideas of Bakhtin,

> *"we consider a voice in a general sense, not reduced to the physical, acoustical dimension. We rather consider it as a distinct, differential position with persistence and interference with other voices. We consider that, for example, an utterance, that means a word, especially if it is repeated ..., an idea, a reply, a book or even a non-verbal act ... may become a distinct voice through its echoes and influences in the subsequent utterances. Of course that we consider as voices also the participants to a conversation or even groups of persons (for example, minorities), because they represent distinct positions, with persistence and that interfere with other voices, be the other persons, groups or voices in a general sense (for example, ideas or replies that influences them)." [1].*

Therefore, in addition to the obvious voices of the participants (explicit or implicit, for example, Newton in our case), we identify voices also starting from the topics discussed in the correspondence. For this purpose and for analyzing the inter-animation of voices, we use and XSLT transformations (https://www.w3.org/Style/XSL/) and the NLP tools developed under the ReaderBench platform [8] of the corpus represented in XML.

An example of using XSLT transformations for the visualization of the distribution of the main concepts in the correspondence of Leibniz and Clarke is shown below.

===**Leibniz Letter nr. 1**
1 - -1 - **soul**, **matter**, **God**,
2 - -1 - **soul**,
3 - -1 - **space**, **God**, **God**,
4 - -1 - **God**, **God**, , **time**, **time**, **time**, , , **God**, **God**, , **God**,
===**Clarke Letter nr. 1**
5 - 1 - **mathematical**, , **God**, **mathematical**, **matter**,
6 - 2 - **soul** **mathematical**,
7 - 3 - **space**, **God**, **God**, **God**, **space**, **God**, **God**, **God**, **God**, **space**, ,
8 - 4 - , **God**, , **God**, **God**, , **God**, , **God**, **God**,
===**Leibniz Letter nr. 2**
9 - 5 - **mathematical**, **mathematical**, **mathematicians**, , **mathematics**, **time**, **mathematical**, **mathematics**, **mathematics**, **God**, , ,
10 - 5 - **matter**, **space**, **matter**, **matter**, **space**, **matter**, **matter**, **God**, **space**,
11 - 7 - **space**, **God**,
12 - 7 - **soul**, **space**, , **soul**, **soul**, **soul**,
13 - 7 - **God**, **soul**,
14 - 8 - **God**, , , **God**, **God**, **God**, **God**, , , **God**, , **God**,
15 - 8 - **God**, **God**,
16 - 8 - **God**, **God**, , **God**,
17 - 8 - , **God**, **God**, , **God**,
18 - 8 - **God**, **soul**,
19 - 8 - **God**,
20 - 8 - , , **time**, **time**, **God**, , **soul**,
===**Clarke Letter nr. 2**
21 - 9 - **mathematical**, , **matter**, **mathematical**, **mathematical**, **mathematical**, , , **God**, **God**,
22 - 10 - **matter**, **mathematics**, **matter**, , **matter**, **matter**, **God**, **matter**, **matter**, **God**,
23 - 11 - **space**, **God**, **space**,
24 - 12 - **soul**, **soul**, , **soul**, , , **space**, **space**,
25 - 13 - **God**, **soul**, **God**,
26 - 14 -
27 - 14 - **God**, **God**, , , **God**, ,
28 - 16 - **God**, **time**, **God**, **God**, , **God**,
29 - 17 - , , **God**, , **God**,
30 - 18 - **God**,
31 - 19 - **God**, **God**, ,

32 - 20 - **God**, **God**, **God**, **God**, **God**, **soul**, **soul**, **God**, **soul**, **God**, **God**, **soul**, **God**, **God**, **God**,

===**Leibniz Letter nr. 3**
33 - 21 - **mathematical**, **mathematics**,
34 - 21 - **space**,
35 - 21 - **space**, **space**, **space**, **God**, **space**, **God**, ,
36 - 21 - **time**, **space**, **time**, **space**, , **space**, ,
37 - 21 - , **space**, **space**, , **space**, **space**, , **space**, **God**, , , **space**, **space**, **God**,
38 - 21 - **time**, **God**, , **time**, **time**, **time**, **time**, **time**, **time**,
39 - 21 - **God**, **space**, **God**, **God**,
40 - 21 - **God**, **God**, **God**,
41 - 22 - **matter**, **God**, **space**, **matter**, **God**, **matter**, **space**, **matter**, **matter**, **space**, **God**,
42 - 23 -
43 - 24 - **soul**,
44 - 25 - **God**, **soul**, **soul**, **soul**,
45 - 27 - , **God**, , **God**,
46 - 28 - **God**, **God**,
47 - 30 - **God**, **God**,
48 - 31 - **God**, **God**, **God**, ,
49 - 32 - **God**, **God**, , ,
===**Clarke Letter nr. 3**
50 - 33 - **mathematical**, **mathematical**,
51 - 34 - , **God**, **matter**, **space**, **space**, **space**, **space**, **space**, , **space**, **space**, **space**,
52 - 35 - , **space**, **God**, **space**, **God**, , **space**, , **time**, **space**, **space**, **space**, **time**, , **time**,
53 - 36 - **space**, **God**, , **God**, , **space**, **time**,
54 - 37 -
55 - 38 - , , **God**,
56 - 39 -
57 - 40 - , **God**, **God**,
58 - 41 - **matter**, , **matter**, , **God**,
59 - 42 -
60 - 43 - **soul**, **soul**, **soul**,
61 - 44 - **God**, **soul**, **God**, **God**,
62 - 45 -
63 - 46 - , , , ,
64 - 47 - **God**,
65 - 48 - **God**, **God**, **God**, , **God**, **space**, ,
66 - 49 - **God**, **God**, , ,

===Leibniz Letter nr. 4
67 - 51 -
68 - 51 - God,
69 - 51 - God, God, ,
70 - 51 - matter, space,
71 - 51 –
72 - 51 -
73 - 51 - space, , space, space,
74 - 52 - , space, space, space,
75 - 52 - space, space, space, space,
76 - 52 - space, God, God,
77 - 52 - space, space, space, space,
78 - 52 - space,
79 - 53 - God, God,
80 - 53 - space,
81 - 53 - God, , time, time,
82 - 55 - space, time, space, time,
83 - 55 -
84 - 57 - space, God, God,
85 - 57 - God,
86 - 57 - God,
87 - 58 - matter,
88 - 58 - matter, matter, matter, ,
89 - 58 - matter, God, matter, , matter,
90 - 59 -
91 - 59 -
91 - 59 -
92 - 59 - God, God, ,
93 - 59 -
94 - 61 - , , God, God, space, space, God, soul, , soul,
95 - 61 - God, soul, , , God,
96 - 61 - soul, soul, God,
97 - 61 - soul, , God, , God, soul, soul, God,
98 - 61 - God, , , soul,
99 - 61 - soul, God, God, ,
100 - 61 - soul, soul, soul, God,
101 - 61 - soul,
102 - 61 - soul,
103 - 63 - , , ,
104 - 63 - ,
105 - 63 - , God, , , ,
106 - 65 - , space, time, time, time, space, time, God,
107 - 66 -
108 - 66 - ,
109 - 66 - God,
110 - 66 - ,

111 - 66 - , , time, , God, space, , God, space, time, God, God, space, God, matter, matter, space, space, space, matter, space, space, matter, space, , matter, matter, space, space, , space, space, matter, , space,
===Clarke Letter nr. 4
112 - 67 - time, time, time, , God,
113 - 69 - God, matter, matter, , matter, God, God, matter,
114 - 71 - time, space, time, time, God, , God, matter, , God, matter, matter,
115 - 73 - space, space, matter, space, matter, matter, time, matter, matter,
116 - 74 - space, matter, , , , space,
117 - 75 - space, space, space, space, space, God, matter,
118 - 76 - , , God, space, God,
119 - 77 - space, space, space, God, space, , space, space, , matter,
120 - 79 - God, , space, mathematical,
121 - 80 - space, space, time,
122 - 81 - God, matter, space, space, God, God, matter, God, time,
123 - 82 - , time,
124 - 84 - , God, God, God, matter, God,
125 - 85 - God, matter, ,
126 - 86 -
127 - 87 - , matter, matter, God,
128 - 88 - God, ,
129 - 90 -
130 - 94 - , God, soul, God, soul, , soul,
131 - 95 - soul, God,
132 - 96 - soul, soul,
133 - 97 - soul, matter, God, soul,
134 - 98 - , , , God, , , ,
135 - 99 - God, soul,
136 - 101 -
137 - 102 - soul,
138 - 103 - , mathematical,
139 - 104 - , , matter,
140 - 105 - God, ,
141 - 106 - space, time, time, space, God,
142 - 107 -
143 - 108 -
144 - 109 -
145 - 110 -
146 - 111 - , , mathematical, matter, matter, matter, matter, matter, , matter, matter,

3.2 Processing with ReaderBench

A first set of processings of the Leibniz-Clarke correspondence corpus were done using the ReaderBench platform, a multipurpose framework, which uses NLP for performing tasks ranging from concept extraction, computing semantic distances between concepts and documents, identification of voices (in the extended sense of the polyphonic model presented in the previous section) and their inter-influence (personal and social knowledge building, inter-animation, etc.). The set of the most frequent concepts were identified and displayed (Figure 3, and a detail of it in Figure 4) and as a network (Figure 5). A classical NLP approach was used: counting the resulted stems after stop-words elimination and stemming. Part of speech tagging was also performed for the separation between nouns and verbs.
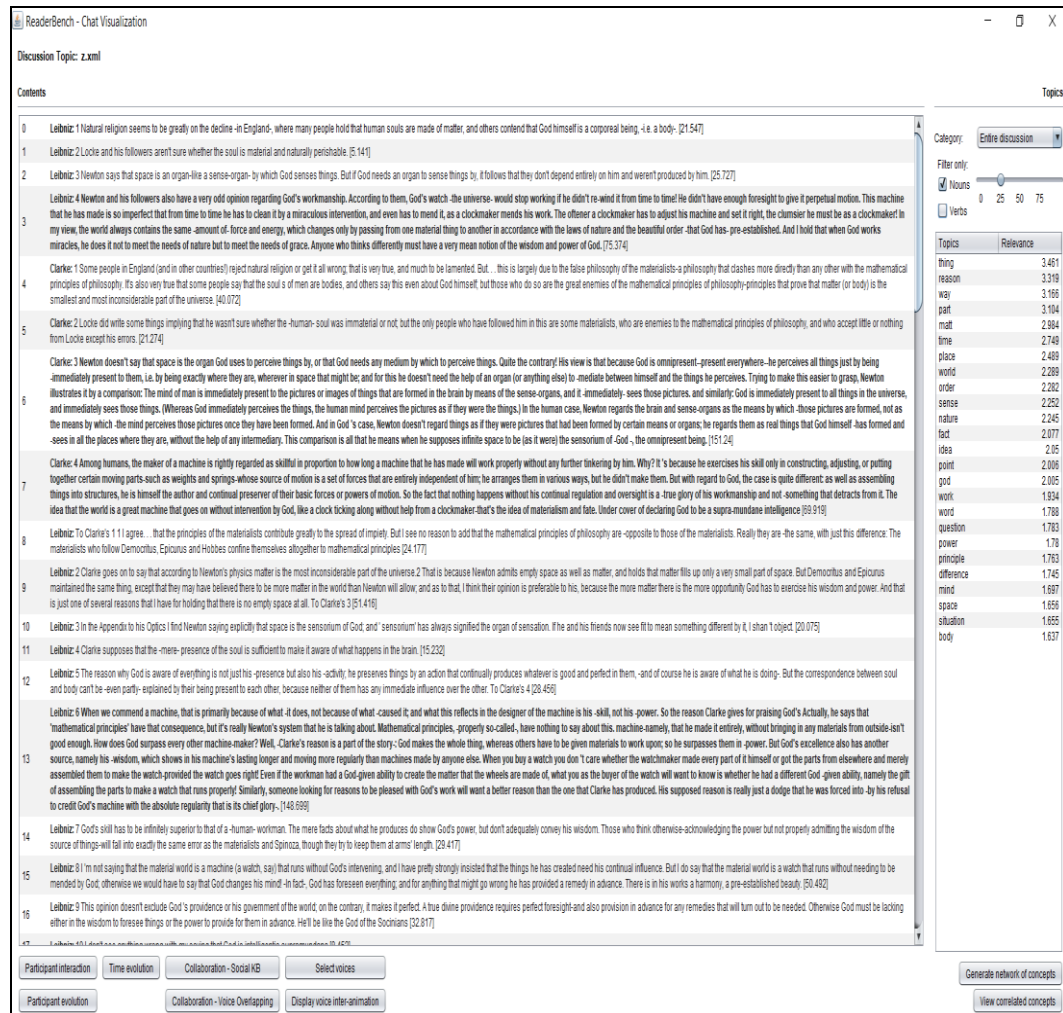
Fig. 3. The basic analysis performed by ReaderBench:
The list of utterances and the most frequent concepts.

Topics	Relevance	
thing	3.461	▲
reason	3.319	
way	3.166	
part	3.104	
matt	2.984	
time	2.749	
place	2.489	
world	2.289	
order	2.282	
sense	2.252	
nature	2.245	
fact	2.077	
idea	2.05	
point	2.006	
god	2.005	
work	1.934	
word	1.788	
question	1.783	
power	1.78	
principle	1.763	
difference	1.745	
mind	1.697	
space	1.656	
situation	1.655	
body	1.637	
cause	1.635	▼

Fig. 4. A detail from Figure 3.

The semantic distance between concepts was computed in ReaderBench using Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA) and distances on the implicit network of sysnsets of WordNet [8]. In Figure 5 the semantic distances between concepts are depicted in the network as physical (Euclidian) distances. The size of the words for concepts reflects the frequence of their apparition in texts.
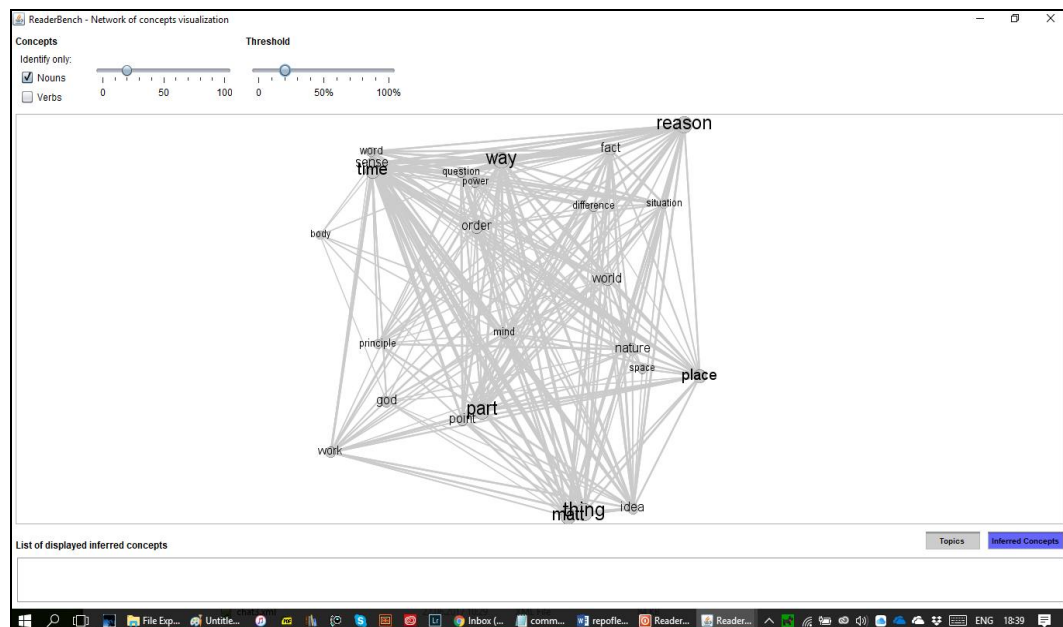
Fig. 5. The semantic distances between concepts.

There were numerous references made by Leibniz and Clarke between the letters, as shown in Table 1. Figure 6 shows their sequence in time as they were displayed by the ReaderBench platform. The length of the rectangles is proportional with the length of the letters, and the relative position of the links correspond to the positioning of the references in text.

Table 1. The number of references

References	Letter number			
	1	2	3	4
Clarke	4	12	17	35
Leibniz	-	12	17	46

ReaderBench uses the LDA method [9] for identifying voices in texts. LDA detects topics, that means sets of semantically close words that are frequently appearing in text. For example, Figure 7 displays the voices identified in the correspondence between Leibniz and Clarke. User may select from the ReaderBench interface the voices that s/he wants to analyze.
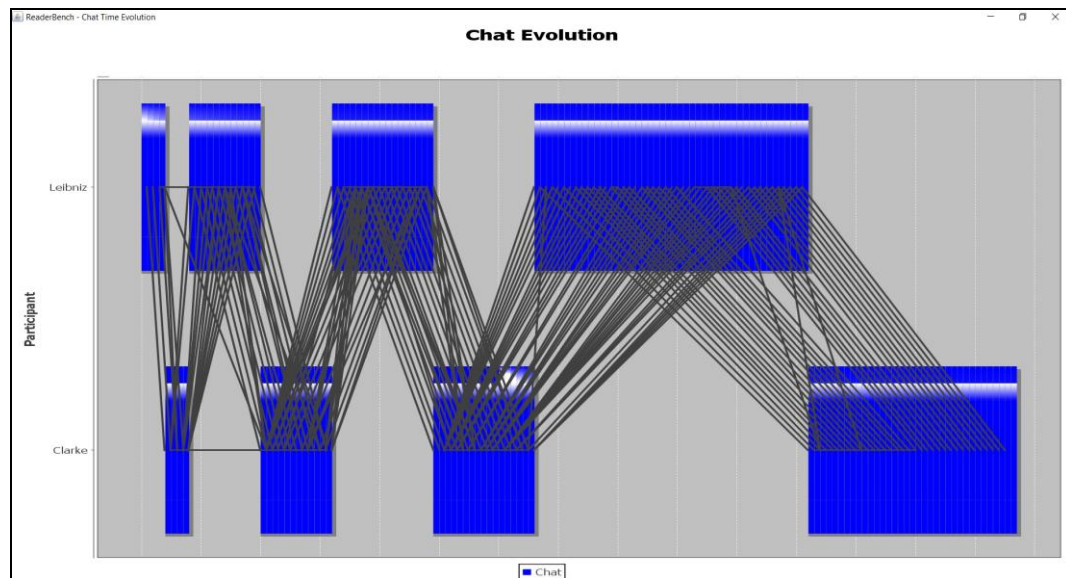
Fig. 6. References between utterances.

For example, the following voices were selected and a statistic of them is displayed in Figure 8:

(god, thing, space)	(sensorium)
(matt) – matter	(word, intelligence, discussion)
(eternal, perfect, perpetual)	(mathematical)
(word, intelligence, discussion)	(sensorium)
(mathematical)	(philosophy, metaphysics)

Fig. 7. Selection of voices.

Fig. 8. Statistics of voices.

ReaderBench displays the inter-animation of voices in Figure 9.

The utterances that are constituted as a voice starting from one topic, are displayed on an horizontal line.

The colors (red and blue, in our case) correspond to the two authors (the "voices" of Leibniz, respectively Clarke).

Some other visualisations are provided for helping the investigation of the joint knowledge construction and the inter-influence of voices (see Figure 10).
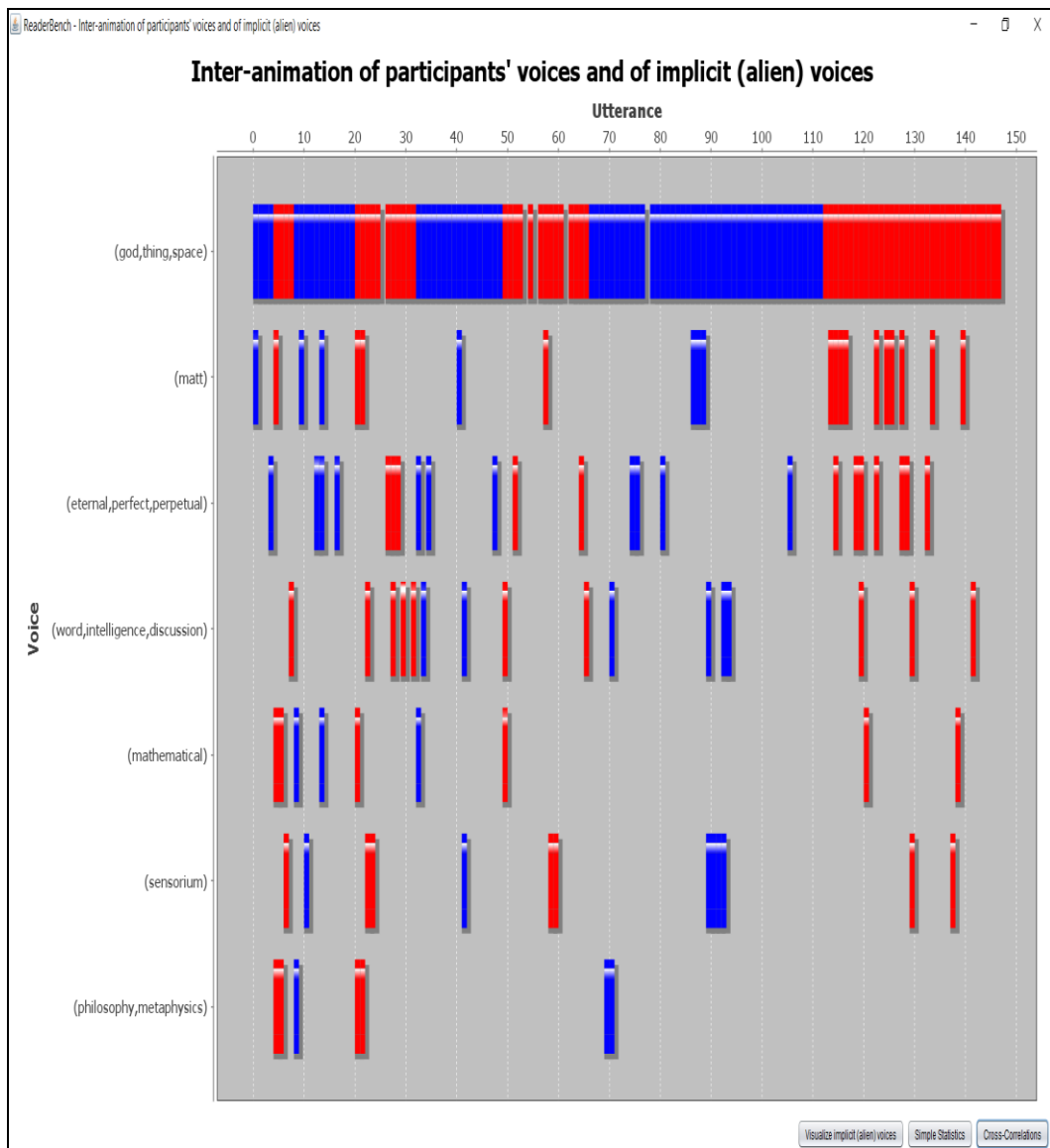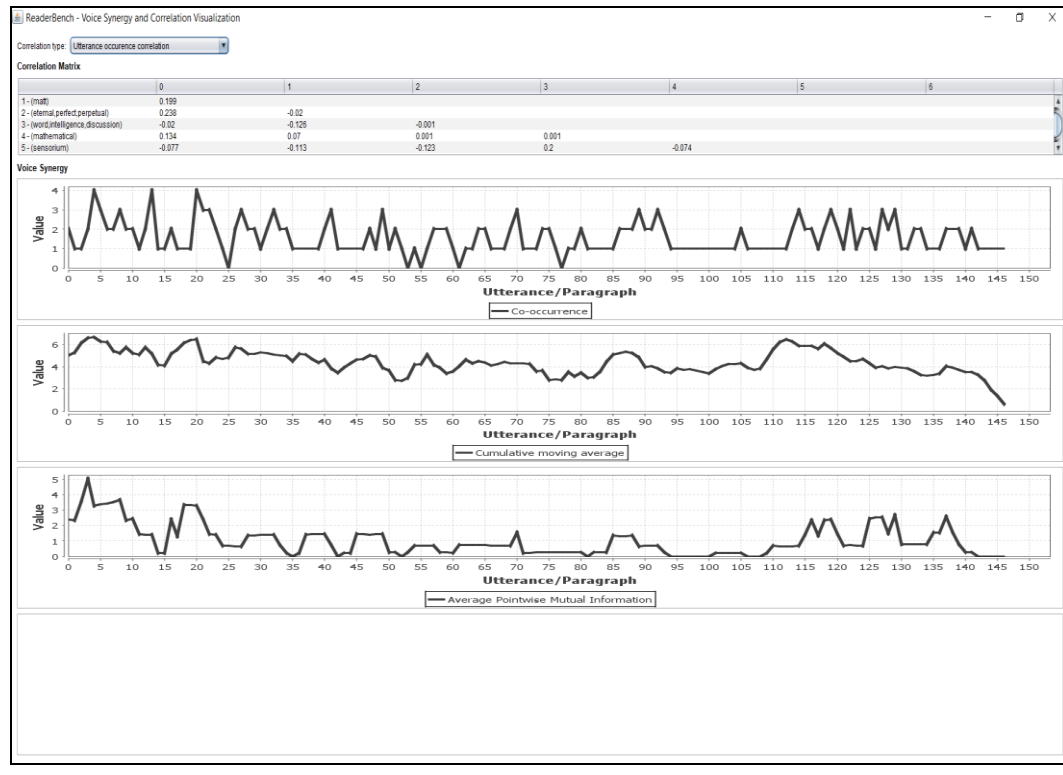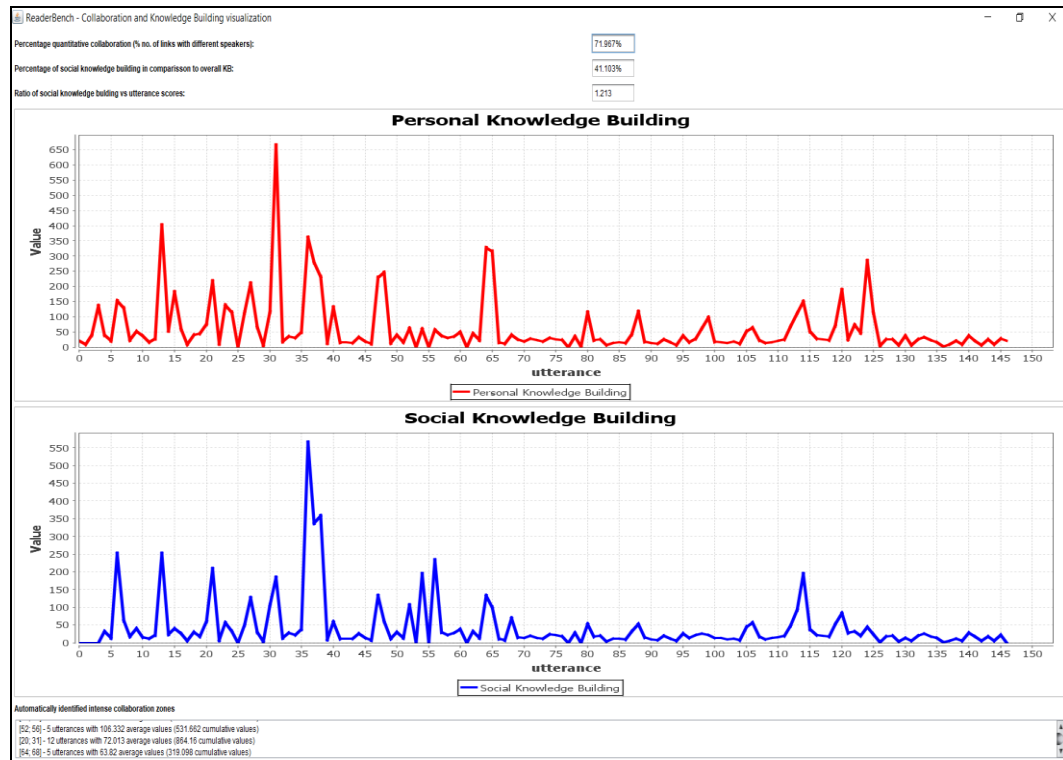

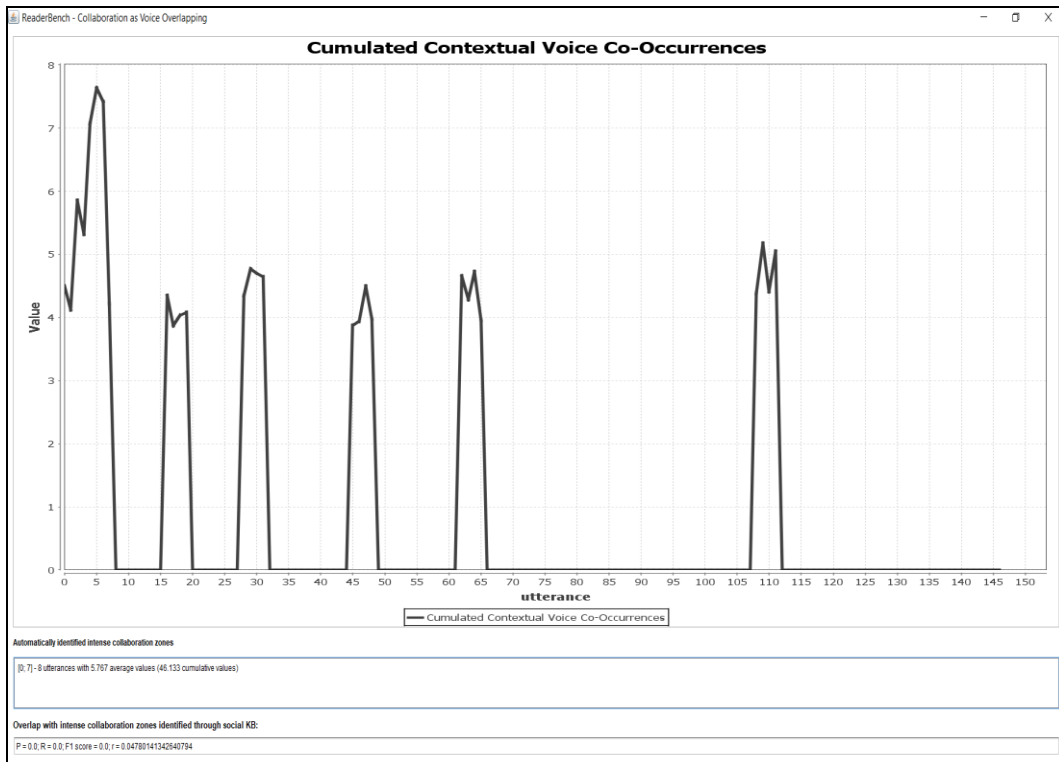**Fig. 9.** Voices' inter-animations.

Fig. 10. Various visualizations of the knowledge construction and voice inter-influences.

4. Conclusions and future developments

Corpora of correspondence containing debates may be annotated, starting from the polyphonic model, in a complex way, which allows the extraction and visualization of various data, such as what are the most important discussed concepts and their semantic relationships, the evolution of the debate, and the inter-animation of the voices (in an extended sense, discussed in the paper).

The annotation structure will be extended for the inclusion of divergences and convergences [7] and for more details for the indirect utterances ("iutt"). Visualization means for these new elements are also under consideration.

One important goal is to enhance researchers to have a hermenophore attitude [10], that means for facilitating a hermeneutic analysis of the texts.

R E F E R E N C E S

[1] Trausan-Matu, S. (**2013**) A Polyphonic Model, Analysis Method and Computer Support Tools for the Analysis of Socially-Built Discourse, Romanian Journal of Information Science and Technology 2-3 (16), Romanian Academy, pp.144-154

[2] Bakhtin, M.M. (**1981**) The dialogic imagination: Four essays (C. Emerson & M. Holquist, Trans.). Austin and London: The University of Texas Press

[3] Vailati, E., (2000) Leibniz şi Clarke. Corespondenţa filosofică, Editura Tehnică, Bucureşti.

[4] ***, A collection of papers, which passed between the late learned Mr. Leibnitz, and Dr. Clarke, in the years 1715 and 1716. Relating to the principles of natural philosophy and religion. With an appendix. To which are added, Letters to Dr. Clarke In the Years 1715 and 1716. Relating to the Principles of Natural Philosophy and Religion. With an Appendix. To which are added, Letters to Dr. Clarke concerning Liberty and Necessity; from a gentleman of the university of Cambridge: With the Doctor's Answers to them. Also Remarks upon a book, entituled, A philosophical enquiry concerning human liberty. By Samuel Clarke, D. D. Rector of St. James's Westminster. Printed for James Knapton, at the Crown in St. Paul's Church-Yard, 1717. Last retrievel at 19 October **2017** from https://books.google.ro/books/about/A_Collection_of_Papers_which_Passed_Betw.html?id =_RUHAAAAQAAJ&redir_esc=y

[5] Trausan-Matu, S., Dascalu, M., Rebedea, T., Gartner, A., (**2010**) Corpus de conversaţii multi-participant şi editor pentru adnotarea lui, Revista Română de Interacţiune Om- Calculator 3(1), pp. 53-64

[6] Trausan-Matu, S. (**2010**) The Polyphonic Model of Hybrid and Collaborative Learning. In F. Wang, L., J. Fong. & R. C. Kwan (Eds.), Handbook of Research on Hybrid Learning Models: Advanced Tools, 2010

[7] Stefan Trausan-Matu, Stahl, G., & Sarmiento, J. (**2007**). Supporting polyphonic collaborative learning. Indiana University Press, E-service Journal, 6(1), 58–74.

[8] Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., Nardy, A. (**2014**) Mining texts, learner productions and strategies with ReaderBench. In: Peña-Ayala, A. (ed.) Educational Data Mining: Applications and Trends, pp. 345–377. Springer, Cham, Switzerland

[9] Blei, D.M., Ng, A.Y., Jordan, M.I. (**2003**) Latent Dirichlet Allocation. Journal of Machine Learning Research, 3(4-5), 993–1022

[10] Trausan-Matu, S. (**2017**) Hermenophore tools, a new perspective on text analysis. *Revista Romana de Interactiune Om-Calculator 10(1)*, pp. 75-88.