**Summarization by Means of the Lexical Cloning Method.**

**Abstract.**

It is proposed that a text is the sum of a subtext and a context. Moreover, the context carries with it a cognitive load, displaying the knowledge that the subject who created the text has of the subjects-and-objects that are dealt with in the text. The subtext, meanwhile, carries with it an emotional load, demonstrating the interrelation between the subject who created the text and

1.  the surrounding world in general, in an abstract sense;

2.  and the subjects-and-objects of the given text in particular.

In this article I affirm that a text has both Aesthetic and Ethical components, where the context provides the Aesthetic component and the subtext the Ethical component.

From a strictly technical point of view, the context and subtext of a text are the summary of its predicative definitions, each of which occurs in the sentences of the text and in the text as a whole with a frequency that is characteristic of the subject who created the text.

Thus the summarization of a text in digest form by means of the Lexical Cloning method consists of the following:

1.  The creation of a summary of the predicative definitions of the subject requiring the creation of the text digest.

2.  The creation of a summary of the predicative definitions of all the sentences that compose the text.

3.     The bringing out of those sentences that have the greatest sum of predicative definitions similar

to the predicative definitions of the question and of the summary of the subject requiring the

creation of the text digest.

That is to say, it is supposed that personalization is indispensable in the creation of a text digest: each

person needs a strictly individual text digest, since his/her need for such is egoistic.

**Foreword.**

Until now, all attempts at the computerized summarization of texts in digest[1] form have been

unsatisfactory, due to a lack of understanding of the true nature of the context of texts, and a simple lack of

awareness of the fact that texts have subtexts. As a result, no method for a truly automatic production of

digests has ever been created, since previous attempts applied the faulty practice of operating by key words

and arbitrarily selected phrases.

---

1 I'm aware of the fact that, by calling a summary of predicative definitions a "summary", I'm creating a

major terminological confusion. But what else should I call a collection of predicative definitions? Out of

necessity I therefore call what are usually called text summaries "digests", and collections of predicative

definitions "summaries", in order to avoid terminological complications in the future; since the Merriam-

Webster Online Dictionary affirms that a digest is "a summation or condensation of a body of information: as

as **a :** a systematic compilation of legal rules, statutes, or decisions **b :** a periodical devoted to condensed

versions of previously published articles".

Indeed, one can't be in any way satisfied with systems that create text digests within an arbitrarily limited volume of texts that have been pre-annotated or supplied with abstracts: the annotation of texts and their abstracts is a very expensive, drawn-out and ineffective process [10,11,12]. Furthermore, human intrusion into the work of the computer is in itself unnatural, since it nullifies the very notion of an automatic processing of the text.

One must note that the problem of the personalization of digests (as results of text summarization) has already come into public view and received some discussion. Unfortunately, there is still the tendency to suppose, as before, that users can formulate their search and summarizing requests by filling out such a form, simply stating what they know and what they are missing in a given situation. In other words, the automatization aspect of formulating what the user needs isn't even mentioned. [1,3,8,10,16]

**CAST.**

One would like to draw attention to the CAST program for creating digests[2]. In my view, the program has the following defect: it doesn't take into account the existence of a text's subtext, and there is no understanding of the nature of context. Moreover, according to the words of the program's creators, the technology of the program contains no element of personalization. [4]

**Borrowings and Adaptations.**

First of all, I would like to clarify the meanings of the terms "predicative" and "non-predicative". As is well known, *praedicatum* in Late Latin means "what has been said (previously)". In Aristotelian and subsequent forms of traditional logic a predicate was understood to be one (the one in which something is said about the

---

2 http://clg.wlv.ac.uk/projects/CAST/programs/VisualisationTool/visualisation.php#contact

subject of speech) of the two terms for the judgment of a subject. In his treatment I count as predicative any

definition of a subject or object in which something is said about an observable subject or object as it

changes.

More: as is well known, Bertrand Russell introduced the notion of a "non-predicative" definition, in which

what is to be defined is brought in through its relation to a class of which it is an element. For example: "the

set of all sets that are not elements of themselves". It is said that the use of "non-predicative" definitions

leads to paradoxes, so they should be dealt with carefully [31,32,33].

**"Kernel Sentences" as Predicative Definitions.**

It has been suggested previously that we refer to minimal constructions carrying a basic load of

meaning in texts as "kernel sentences" [23]. It has also been supposed that kernel sentences must provide

an unambiguous, formal description arising from the subjects/objects-and-their-activity of the given Names[3].

But I'm a philosopher. In this case, it seems more expedient to use a philosophical terminology -- the

terminology used by Aristotle, Moore, Russell and Wittgenstein, as the most significant philosophers to have

dealt closely with the study of language. Those philosophers spoke of predicative and non-predicative

definitions for the description of events and changes occurring in subjects or objects of the Universe.

Moreover, predicative definitions and "kernel sentences" are (almost) one and the same thing.

---

3 "...to understand a sentence it is necessary to know the kernel sentences from which it originates".

Chomsky, Noam. *Syntactic Structures.* The Hague: Mouton, *1957.*

The technology for extracting kernel sentences from a text – as well as the method for formalizing "kernel sentences" – did not exist previously and was created by myself. A more detailed description of this process can be found in US Patent 6.199.067.

To a certain extent one can affirm that it is predicative definitions that create "continuity of lexical meaning" [34], although it may be more correct to speak of the extension, the continued admissibility of the real meaning of a text from sentence to sentence, from paragraph to paragraph, etc.

**The Text.**

A text is made up of words. But what is a word? First of all, a word is made up of letters, which are, in practice, meaningless if separate. And yet a word, as the joining together of several letters, already, beyond any doubt, has a certain meaning. But the existence of synonymy makes the meaning of words, taken separately, vague and lacking in concreteness and separate words are declared to be non-predicative definitions. For example, the word "red", taken by itself, can mean anything: beginning with a colour and ending with a pejorative name for a Communist. In order to understand the "true" meaning of a word one must first identify in what minimal lexical construction of speech a given word is being used; where a minimal lexical construction of speech is a predicative definition: the articulation of three words, relating to three parts of speech - substantives, verbs and adjectives - in the context of a sentence, paragraph or text. All other parts of speech, with the exception of prepositions and interjections, can be, in some way or other, taken to be (or be reduced to) nouns, verbs, and adjectives, where:

1. A substantive has the meaning of the abstract Name of certain objects and subjects;

2. A verb defines the abstract Name of an action;

3. An adjective is the abstract Name describing objects and subjects in the process of change.

This triad of non-predicative definitions is indispensable to defining the subjective evaluation of a fact, when faced with the possibility and the need to include the objects and subjects of the fact within the certain context and subtext. In other words, a person has to evaluate a sandwich from all sides: he has to understand that it's a sandwich, and to decide whose it is, whether he should eat it or not, whether it's fresh and tasty, etc.

The presence of at least one predicative definition is absolutely necessary and sufficient for the creation of a sentence, even if it's missing one or more words from the substantive/verb/adjective triad. Such a word or words can be reconstructed on the basis of the context and subtext of the text; where:

1. The context consists of those predicative definitions where a substantive is used as the abstract Name of objects and subjects and abstractions;

2. The subtext consists of those predicative definitions where pronouns and interjections are used as the abstract Name of objects and subjects and abstractions.

For example, having said the word "unfresh", one can reconstruct the words "sandwich" and "exists" if we know in what context and subtext the word "unfresh" appears. And if we don't know the context and subtext of a given predicative definition, then the word "unfresh" can be used with, for example, the words "fish" and "smells". Only a text, being a collection of predicative definitions grouped together in meaningful sentences, can provide, more or less identically, the context and subtext of every one of these predicative definitions. That is, a text is considered to be completed in so far as its context and subject are, more or less identically, defined.

**Processing the Text.**

The task comes down to extracting all the predicative definitions from every sentence of the text. Such a collection is termed a summary of a sentence; the number of times each predicative definition occurs in the sentence, paragraph or text is referred to as its weight. (NLP uses the notion of predicative definitions not counting weights of them.) A summary, being an ordered list of predicative definitions, is susceptible to rapid processing by computer.

In principle, the extraction of predicative definitions could be called the "decomposition" of the text, but "decomposition" only for the purpose of creating a summary of the text's predicative definitions. Thus in preparation for DUC 2004 – in which for various reasons I couldn't take part – the question arose: when creating super-short digests, shouldn't those digests consist of predicative definitions rather than of pieces of sentences from the texts?

**Examples of Summary.**

At the URL http://lexiclone.com/SummarySample_Fyodor_Dostoevsky.htm ,

the reader can see an extract from Fyodor Dostoevsky's summary (a part of which is reproduced below), created on the basis of his book *The Brothers Karamazov* in English (the numbers to the right represent the frequency of each predicative definition's occurrence in the text):

it - be - in : 1 466

i - be - in : 1 347

it - have - in : 996

you - be - in : 936

you - be - your : 798

i - have - in : 664

all - be - in : 657

it - will - in : 535

my - be - in : 496

all - have - in : 473

Clearly, it is subtext that dominates in Dostoevsky's text. In another summary – that of my Patent

#6.199.067 – it is context that dominates:

one - say - least : 1 447

segment - say - least : 1 124

datum - item - plural : 1 025

system - say - remote : 950

datum - say - plural : 888

computer - say - remote : 845

datum - item - linguistic : 845

system - say - least : 844

computer - say - least : 818

one - say - remote : 805

**The Ethics and Aesthetics of a Text.**

Thus it is supposed and demonstrated (cf. the fragments of summaries above) that a text is the integrated sum of a subtext and a context. Moreover, the context carries with it a cognitive load, displaying the knowledge that the subject who created the text has of the subjects-and-objects that are dealt with in the text. The subtext, meanwhile, carries with it an emotional load, demonstrating the interrelation of the subject who created the text with the surrounding world in general, and with the subjects-and-objects of the given text in particular. Indeed, the function of pronouns is to abstract from the Names that are signified by means of substantives and other parts of speech equated with them.

I go on to affirm that the subtext is the Ethical component of a text: Ethics relates to the egoistic evaluation of the consequences that would result from the relationship of the subject who created the text with the surrounding world, abstracted from concrete subjects-and-objects.

The Aesthetic component, meanwhile, is the context of a text: the knowledge of the concrete subjects-and-objects of a concrete text.

**Cynicism.**

I belong to the Cynical school of philosophy, the foundations of which were laid by those contemporaries of Plato and Aristotle, Ecclesiastes and Jeremiah. Those philosophers – Plato and Aristotle as well as Ecclesiastes and Jeremiah – affirmed with one voice that something exists – which they called "virtue" – which forces the world to change. Without going into detail, I, as a Cynic, consider this "something", this "virtue" of a person (as far as the theme of this article is concerned), to be egoism[4]. That is

---

4 Merriam-Webster Online Dictionary assumes:

to say, I'm certain that every subject requires that digest of the text which he needs egoistically, due to egoistic considerations.

**Demonstration of the Underlying Egoism of the Ethical and Aesthetic Components of a Text.**

What gives me the right to affirm that the subject needs a text digest due to egoistic considerations?

Cynicism looks at the whole world through the prism of the Theory of Internal Relations: everything that happens in the Universe is always being interpreted through changes in the internal constitution of a subject and can only be evaluated subjectively. For this purpose one uses the preposition "in", which always and simultaneously corresponds to the adjective "interior" -- in all human languages, without exception! Indeed, the remarkably frequent use of the adjective "interior" in the most commonly encountered predicative definitions in all languages leads one to the following thought: in the structure of predicative definitions the adjective/preposition "in-interior" indicates the subject's familiarity with the object/subject of a given predicative definition and its acceptance into the subject's interior world. I equated egoism to the desire to include, to have something "into", "within" one's Uneverse.

---

Main Entry: **ego·ism**

Pronunciation: 'E-g&-"wi-z&m *also* 'e-

Function: *noun*

**1 a :** a doctrine that individual self-interest is the actual motive of all conscious action **b :** a doctrine that individual self-interest is the valid end of all actions

**2 :** excessive concern for oneself with or without exaggerated feelings of self-importance

For an obvious demonstration of this speculation it will be enough to glance at the summary excerpts quoted above. In addition, we present below a superficial, not altogether authentic analysis of a few summaries, which are also to be taken as demonstrations of the abovementioned theory about the egoistic nature of human "virtue", as that of the striving to have everything "into".

To take an example, in the divine Julius Gaius Caesar's summary, created on the basis of his work *De Bello Gallico*, we can see that the most frequently occurring predicative definition, which has led to the very word "Caesar" becoming a pejorative, is the triad "caesar-be-in'. The next triads in the series, such as "all-be-in", "all-have-in", and "caesar-have-in" are understood as an attempt to include absolutely everything "inside". The army – "army-be-in" the legions – "legion-be-in" and the camps – "camp-be-in" – all have to be there, too. And all the divine one's enemies – "enemy-be-in" and "gaul-be-in" – must follow "into" there as well. One can assume that it is precisely the predicative definitions containing the adjective "interior"/ preposition "in" that reflect Julius' acknowledgment of the subjects/objects of the predicative definitions.

Speaking of Daniel Defoe, we also see that everything must be "in" the Universe known by Daniel Defoe – "all-be-in". There's the Cynical question – "if-be-in?" to which Defoe himself gives the answer – "much-be-in". This is my reading of Robinson Crusoe.

And now, the radiant Anton Chekhov. He believes "in" God "in" the Universe – "one-be-in" – and that we will all be there – "all-be-in". At the same time Anton Chekhov puts to himself the Skeptical (the same, Cynical) question: "if-be-in"? He is certain that even there, in that bright world, lies will penetrate – "lie-have-in". Nevertheless, he is ready to take everything into the Universe – "out-be-in". After all, his Universe is full

of love – "love-be-in". But falsehood will penetrate even there – "lie-will-in", - where nothing is "nothing-be-in". This is my reading of "Kashtanka" and other short stories .

And here are the ideas that predominated "in" the mind of Lenin "in" the year 1919: the country, the Party, the Communists, the bourgeoisie, and all people must be in the whole – "country-be-in", "party-be-in", "all-be-in", "communist-be-in", "bourgeois-be-in", "mass-be-in". Everything, "inside" his Universe, must be Soviet – "soviet-be-in"; and all classes (social classes, obviously) must be in the whole – "class-be-in". As a true revolutionary – "revolution-be-in" - Lenin is persuaded that struggle is everything – "struggle-be-in". In such conditions the Party must be the party of the majority, the party of the Bolsheviks, and not just anything – "party-be-party", "must-be-in", "most-be-in", "within" his Universe. Nevertheless, Lenin is a true believer: "one-be-in", although the predicative definition "party-be-in" has a significantly greater weight[5].

The presence in all human languages, without exception, of the preposition "in", which always and simultaneously corresponds to the adjective "interior" -- and the remarkably frequent use of it in predicative definitions constitutes the ultimate verification of Cynical Skepticism, a factual confirmation of the philosophy of Cynicism and of the egoistical nature of one's need for summarization!

**Cynicism: Time and Eternity.**

At the basis of Cynicism is the idea that there is a world of change in time and a world without change in eternity. In other words, everything that exists changes in time. And what isn't in time is unchanging in eternity.

1. Reality.

---

5 Samples of summaries are at the site http://lexiclone.com

Reality is what doesn't exist in time but exists in the eternity of immutability. This means that everything that exists in completed time – in the past, in the present, or even the future if it is in completed time – is Reality.

2. Truth.

Subjective opinions, however, are Truths; and subjective Truths are always opinions (5). Indeed, opinions change and the search for Truth is always subjective6.

**Lexical Cloning.**

Lexical cloning is the computerized indexing of the predicative definitions in the texts of a particular author, calculating the frequency with which given predicative definitions occur in the text. In the beginning lexical cloning was thought to be a simple demonstration of the possibilities of an information-search technology.

**Synonyms as the Mechanism Linking Predicative Definitions for "Continuity of Lexical Meaning".**

Synonymic expansion of predicative definitions (at a purely technical level) means that all possible synonyms are found for every word in every predicative definition, and every possible combination of these words is structured according to the substantive-verb-adjective pattern. The term "by association" indicates that the subjects and/or objects of an event resemble (or the event itself resembles) -- in some cases, very distantly – the subjects and/or objects of events described by other predicative definitions.

---

6 "For we know in part and we prophesy in part. But when that which is perfect has come, then that which is in part will be done away." Paul. Bible. I to Corinthians, 9&10

That is to say, "continuity of lexical meaning" is realized by the use in the text of a set of synonymous predicative definitions, and *not* of non-predicative definitions, as in the case of words standing by themselves.

**Summarization.**

Therefore, if one considers my previously enumerated speculations to have been demonstrated – namely:

1.  that texts have an Ethical and an Aesthetic component,

2.  that texts carry in themselves information about Truth and Reality,

3.  that the basic orientation of the subject in the creation of a text digest is always egoistic,

then the question arises: which sentences must be picked out of the text in the digests that result from summarization?

I affirm that:

a)  We have a bipolar system: there is on the one hand a text created by a subject, and there is on the other hand a subject who for egoistic reasons needs the digest of that text. It's obvious that the Ethical and Aesthetic conditions of the two subjects are likely to be more or less different. In other words, the two subjects are likely to have different knowledge of the subjects-and-objects of the text, and different ways of conceiving them and the surrounding world. The task, then, comes down to how one should select those sentences from the text of one subject that are egoistically – Ethically and Aesthetically – appropriate to the other subject. In other words, one must select in a

text sentences which, on an emotional and cognitive level, egoistically – and thus strictly

subjectively – are appropriate to one.

b) Reality has the nature of singleness: when a subject asks for information about Reality, it's only

possible to find one single and monosemantic Reality in another subject's text. Indeed,

1. if a subject is interested in the digest of some text that must contain information about the

date of Christ's birth,

2. and that text contains information about the date of Christmas according to the Gregorian

calendar,

3. he will inescapably be informed of the Reality of the date of Christmas according to the

Gregorian calendar. And all other subjects will receive exactly the same digests,

mentioning the date of Christmas according to the Gregorian calendar.

c) Truth, meanwhile, is subjective, and one subject egoistically seeks in another's text that Truth

which is specifically and exclusively (egoistically) appropriate to him. That is to say, if one is

seeking the Truth about how to get from the Metropolitan Museum to Grand Central, one could be

presented with several Truths to choose from:

1. one could take a taxi,

2. one could take the subway,

3. one could buy a car,

4. one could hire a rickshaw,

5. one could go on foot.

**The Creation of the Summary of the Predicative Definitions of the Subject Requiring the Creation of a Text Digest.**

It follows from what has been said above that it is necessary to create a summary of the predicative definitions of the subject who requires the creation of a text digest, that it is important to know what one needs.

Such a summary of predicative definitions must inevitably contain those predicative definitions that are most habitual for the subject requiring the creation of the text digest -- those that abundantly characterize his egoistic requirements of the text digest from both the Ethical and the Aesthetic standpoint.

That is to say, such a summary must contain information about which specific method of getting to Grand Central would be optimal for the subject who needs the text digest.

**Compatibility.**

The closeness between the digest to one's needs is defined according to a standard formula, called Compatibility:

$$\text{Compatibility} = \left( \frac{\text{Sum} \left( \text{Weight-of-the-same-predicative-definition-User} \ * \ \text{Weight-of-the -same-predicative-definition-Text} \right)}{\text{sqrt} \left( \text{Sum} \left( \text{Weight-of-each-predicative-definition-User}^2 \right) \ * \ \text{Sum} \left( \text{Weight-of-each-predicative-definition-Text}^2 \right) \right)} \right) * 100$$

**Tautology.**

Compatibility arrives at maximum significance -- at 1 (100%) -- only in the case where the search results in a tautology -- only when what was asked is repeated literally.

This means that Ludwig Wittgenstein was right: ". the tautology and the contradiction that they say nothing. The tautology has no truth-conditions, for it is unconditionally true. Tautology and contradiction are, however, not nonsensical; they are part of the symbolism, in the same way that "0" is part of the symbolism of arithmetic[7]."

In other words, with the summarization of a text into a digest – with the application of the forementioned Compatibility to the search for sentences – only those Truths of one subject will be included that stand closest to the subjective, egoistic Truth of the other. The only exceptions – in the extreme case of complete tautologies (completely identical sentences) being included in the digest – would be agreement with (or contradiction of) one subject's Truth by another.

For example, a rich eccentric's digest could include the Truth of buying a car or hiring a rickshaw as means of getting from the Metropolitan Museum to Grand Central.

Realities in digests, meanwhile, will always be tautological in the sense that the answer to a question will always be a monosemantic answer independent of the nature of the one asking.

**Just a Digest.**

It's possible to create a text digest without taking into account the needs of the one asking for the digest. In that case what takes place is the extraction of sentences from the text that carry a certain informational load and an Aesthetic-Ethical filling, corresponding to the nature of the author of the text. Evidently, that CAST creates such digests calling them "summaries".

**Conclusions.**

---

7 [4.4] Ludwig Wittgenstein. Tractatus Logico-Philosophicus. http://www.kfs.org/~jonathan/witt/ten.html

The summarization of a text thus comes down to finding those sentences that have the greatest sum of weights of predicative definitions that correlate to the ones in the summary of the subject who is in egoistic need of the digest.

In other words, every subject needs an individual approach to the creation of a text digest: the text digest must reflect the Ethical and Aesthetic predilections of the subject and contain the Truths and Realities he seeks.

If one does not personalize the process of creating a digest and uses CAST, the digest of a text will only reflect the Ethical and Aesthetic peculiarities of the text's author.

**References**

1. **Becher, Margit; Endres-Niggemeyer, Brigitte; Fichtner, Gerrit: Scenario forms for web information seeking and summarizing in bone marrow transplantation Coling Workshop for multilingual summarization and question answering, Taipeh, 2002**

2. **Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. 1995. Centering: A framework for modelling the local coherence of discourse.*Computational Linguistics,* 21(2):203 – 225.**

3. **Eugene Charniak. *Taggers for parsers*, (with Glenn Carroll, John Adcock, Antony Cassandra, Yoshihiko Gotoh, Jeremy Katz, Michael Littman, and John McCann), Artificial Intelligence (forthcoming).**

4. **Constantin Orasan, Ruslan Mitkov, Laura Hasler (2003): CAST: a Computer-Aided Summarisation Tool. In *Proceedings of Research Notes Sessions of the 10th Conference of***

*The European Chapter of the Association for Computational Linguistics (EACL2003)*, Budapest, Hungary

5.  Craven T.C. 1996. An experiment in the use of tools for computer-assisted abstracting. In *Proceedings of the ASIS 1996*.

6.  Craven, T.C. 1986. *String Indexing*. Orlando: Academic Press.

7.  Ehud Reiter and Somayajulu Sripada. Human Variation and Lexical Choice. Computational Linguistics. Volume: 28, Number: 4, Page: 545-553

8.  Endres-Niggemeyer, B; Wansorra, E.: Making Cognitive Summarization Agents Work In A Real-World Domain in Sharp, B. ed.: Natural Language Understanding and Cognitive Science. 1st Int. NLUCS Workshop. Porto, Portugal, April 2004, 86-96.

9.  Endres-Niggemeyer B. 1998. *Summarizing information*. Springer.

10. Halliday, M.A.K., and R. Hasan. 1989. Language, Context and Text: Aspects of language in a social-semiotic perspective. Geelong, Victoria: Deakin University Press. (republished by Oxford University Press 1989).

11. Halliday M., Hasan R. 1976. Cohesion in English. Longman, London.

12. Hongyan Jing. Using Hidden Markov Modeling to Decompose Human-Written Summaries. Computational Linguistics. Volume: 28, Number: 4, Page: 527-543

13. Horacio Saggion and Guy Lapalme. Generating Indicative-Informative Summaries with SumUM. Computational Linguistics. Volume: 28, Number: 4, Page: 497-526

14. Ilya S Geller. 2003 The Role and Meaning of Predicative and Non-predicative Definitions in the Search for Information· NIST TREC 2003

15. Janyce Wiebe ; Theresa Wilson ; Rebecca Bruce ; Matthew Bell ; Melanie Martin. Learning Subjective Language. Computational Linguistics. Volume: 30 Number: 3 Page: 277 -- 308

16. Jane Morris, Graeme Hirst. The Subjectivity of Lexical Cohesion in Text. http://ftp.cs.toronto.edu/pub/gh/Morris+Hirst-2004-EAAT.pdf

17. Laura Hasler, Constantin Or¢asan, and Ruslan Mitkov. 2003. Building beter corpora for summarisation. In *Proceedings of the Corpus Linguistics Conference*, Lancaster, UK, 28th – 31th March.

18. Mann, William C. and Thomson, Sandra A. (1988). Rhetorical structure theory. *Text*, 8, 243– 281.

19. Masumi Narita. 2000. Constructing a tagged EJparallel corpus for assisting Japanese software engineers in writing English abstracts. In *Proceedings of the Second International Conference on Language Resources and Evaluation*, pages 1187 – 1191, Athens, Greece, 31 May – 2 June.

20. Michael Hoey. 1991. *Patterns of Lexis in Text.* Describing English Language. Oxford UniversityPress.

21. Mitkov R. 1995. A breakthrough in automaticabstracting: the corpus-based approach. Technical report, University of Wolverhampton.

22. Morris, J., and Hirst, G. (1991). Lexical cohesion computed by thesaural relations as an indicator of the structure of text. Computational Linguistics, 17(1), 21–48.

23. Noam Chomsky. 1957. *Syntactic Structures*.

24. Klaus Zechner. 2003. Automatic Summarization of Open-Domain Multiparty Dialogues in Diverse Genres. Computational Linguistics. Vol. 28, Issue 4. p. 447 - 485

25. Klaus Zechner. 1996. Fast generation of abstracts from general domain text corpora by extracting relevant sentences. In COLING – 96

26. Rebecca Hwa. Sample Selection for Statistical Parsing. Computational Linguistics, Vol. 30, Issue 3 - September 2004, pp. 253 - 276

27. Robert L. Donaway, KevinW. Drummey, and Laura A. Mather. 2000. A comparison of rankings produced by summarization evaluation measures. In Proceedings of NAACL-ANLP 2000 Workshop on Text Summarisation, pages 69 – 78, Seattle, Washington, April 30.

28. Sanmay Das, Barbara Grosz and Avi Pfeffer. 2002. Learning and Decision-Making for Reconciliation. In the Proceedings of the First Joint Conference on Autonomous Agents and Multi-agent Systems, Bologna, Italy. July, 2002.

29. Simone Teufel and Marc Moens. 2003. Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status. Computational Linguistics. Vol. 28, Issue 4. p. 409 - 445

30. Zechner K., 2002. Summarization of Spoken Language - Challenges, Methods, and Prospects. Speech Technology Expert eZine, Issue 6, January 2002.

31. Bertrand Russell, Scientific Method in Philosophy.

32. Bertrand Russell, Introduction to Mathematical Philosophy.

33. Bertrand Russell. My Philosophical Development.

34. Jane Morris, Graeme Hirst. Lexical cohesion computed by thesaural relations as an indicator of the structure of text. Computational Linguistics. Volume 17 , Issue 1  (March 1991)