



**HAL**  
open science

## SENTIMENT ANALYSIS: AUTOMATIC CALCULATIONS IN NOOJ ENVIRONMENT

Francesco Terrone, Ritamaria Bucciarelli, Andrea Rodrigo, Javier Julian  
Enriquez, Francesco Saverio Tortoriello, Ilaria Veronesi

► **To cite this version:**

Francesco Terrone, Ritamaria Bucciarelli, Andrea Rodrigo, Javier Julian Enriquez, Francesco Saverio Tortoriello, et al.. SENTIMENT ANALYSIS: AUTOMATIC CALCULATIONS IN NOOJ ENVIRONMENT. 2024. hal-04478676

**HAL Id: hal-04478676**

**<https://hal.science/hal-04478676v1>**

Preprint submitted on 27 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SENTIMENT ANALYSIS: AUTOMATIC CALCULATIONS IN NOOJ ENVIRONMENT

F. Terrone<sup>1</sup>; R. Bucciarelli<sup>2</sup>; A Rodrigo<sup>3</sup>; J. Julian Enriquez<sup>4</sup>; F.S. Tortoriello<sup>5</sup>; I. Veronesi<sup>5</sup>

<sup>1</sup> *Fondazione F. Terrone. Università degli studi di Napoli Federico II*

<sup>2</sup> *Università di Venezia*

<sup>3</sup> *Universidad Nacional de Rosario [Santa Fe]*

<sup>4</sup> *Universitat Politècnica de València*

<sup>5</sup> *Università degli Studi di Salerno*

## Abstract

Smart mobility has a positive impact on the lives of our cities through new technological solutions. An economy driven by an immense amount of data that allows us to engage in new business. A new cultural perspective and a new education approach where knowledge merges and generates new knowledge osmotically. We explore teaching strategies and excellent skills, which can meet the needs of higher education, such as the latest generation of technologies used in business, social and personal life, as well as in institutions and colleges of higher education.

Technical fields like scientific and distributed informatics, and computational linguistics have new connotations as knowledge is reformulated and disciplines acquire new meanings. The research team aims to study the characteristics of the LSP language communication systems, describe and scientifically validate the characteristics, through conjectures and refutations of assumptions proposed by linguists, computer specialists and technologists. The phases of the research are divided into several chapters: research of textual production analysis - descriptions - comparisons with a focus on new text production techniques. - Second phase of data processing in NLG. Environments and processing techniques; third phase of validation and emotional techniques in the advertising text.

**Keywords:** Scientific computer science, computational sciences, high performance computing, text production techniques, data processing in NLG.

## ANÁLISIS DE LOS SENTIMIENTOS: CÁLCULOS AUTOMÁTICOS EN EL ENTORNO DE NOOJ

### Resumen

*La movilidad inteligente influye en la vida de nuestras ciudades con nuevas soluciones tecnológicas. Una economía impulsada por una inmensa cantidad de datos que nos permite hacer nuevos negocios. Abre un nuevo horizonte cultural y una nueva educación en la que el conocimiento se funde y genera osmóticamente nuevos conocimientos. Exploramos estrategias de enseñanza y excelentes habilidades, que pueden satisfacer las necesidades de la educación superior, como la última generación de tecnologías utilizadas en los negocios, la vida social y personal, así como en instituciones y colegios de educación superior.*

*El conocimiento es reformulado y las disciplinas adquieren nuevas connotaciones especializadas en campos científicos como la informática científica y distribuida y la lingüística computacional. El equipo de investigación tiene como objetivo estudiar las características de los sistemas de*

*comunicación lingüística LSP, describir y validar científicamente las características, a través de conjeturas y refutaciones de suposiciones propuestas por lingüistas, especialistas en informática y tecnología. Las fases de la investigación se dividen en varios capítulos: investigación de análisis de producción textual - descripciones - comparaciones con un enfoque en nuevas técnicas de producción de texto. - Segunda fase del tratamiento de datos en NLG. Entornos y técnicas de procesamiento; tercera fase de validación y técnicas emocionales en el texto publicitario.*

**Palabras clave:** *Informática científica, ciencias computacionales, computación de alto rendimiento, técnicas de producción de texto, procesamiento de datos en NLG.*

## **1. Introduction**

The use of synthetic communication, which includes fixed sentences or paraphrases of fixed sentences (pre-contextualized texts), is becoming more prevalent in the language galaxy. Linguistic resources that can ensure specialized communication are required for this type of communication. That is, models that are precise, but are created using terminologies that are specific to certain professional fields and reproduce a short semantics. The formal grammar and phrasing order are the most noteworthy features (Bucciarelli 2006, p. 24). The research question that the team asks itself is the following: What education and tools must a trainer of excellence use to convey knowledge? Will technical specialist languages completely replace traditional technical skills? Will we be able to overcome the dispute about the codes and the discussion of communication procedures that are perfect, punctual, on the other hand at the borders of the planetarium, and on the other, with obsolete and imprecise definitions?

In this research hypothesis, the team tries to validate the evolution of the ESP languages using linguistic techniques and explain their validations using irrefutable mathematical models such as: 1) Characteristics of the LSP-NLG-ESP languages with a scientific description of the NLG model. 2) Starting from the premise that LSP languages are now classified as ESP, the team scientifically explains their NLG production by proving that professional texts are not emotion-free and can lead to ambiguities in NLG production: will this be true? The pedagogical potential of corpus work was first appreciated by ESP teachers in this context, as per Bernardini (2004) since the corpus approach's advantage is its capacity to merge frequency data for all words in a corpus and the verbal context in which they appear.

## **2. Man builder text: phase A**

The Mood Woman, the virtual director, is the symbol of our society in the intelligent era. In this way, it is logical to start by thinking about what ideologies and values to transmit. If Mood represents the latest technologies, it may only convey actions that are the result of acronyms being absorbed by algorithms in high-computing environments. What semantics are necessary to convey? The mood is influenced by excellence, mind triumph in human actions, and knowledge of yesterday and today, which means that one is wrong.

The research's focus lies in the planning and decision-making of suitable conduct criteria for establishing the data collection model and for selecting the descriptive linguistic methodology lexicon-grammar (Elia *et al.*, 1981), as it provides a description of the combinatorial mechanisms (morphosyntactic, distributive, and transformational) associated with specific lexical input.

It is the phase in which people generate and transform a language for their functions, according to anthropology, sociology, and culture. He/she humanizes and becomes a language for the areas. He/she is the builder of his/her knowledge and uses descriptive techniques, grammars of the language being used, and then establishes criteria and characteristics. To put it simply, they develop a specialized LSP language for intelligent communication that is approved for both fixed and synthetic applications. For data collection during the design phase, they implement some linguistic structures and emphasize syntactic elements. Their identification of writing mechanisms includes the synthetic techniques of minimum sentences and keyword structuring.

No more forecast data, but certain irrefutable data. This is where the research intervenes to validate these textual aspects of the LSP communication, in the context of the topic, with appropriate examples on central themes of the syntax to conjecture or refute the proposed models. Gotti (1991) defines among the various denotations of special language...

The term "special language" could also be reserved for those languages that, although they generally share the communicative conventions of a particular language, also have some that are not included in the common linguistic heritage. One of those languages, for example, is SEASPEAK (Gotti, 1991, pp. 6-7). Hoffman (1984) describes the fundamental criteria that characterize specialized languages as accuracy, simplicity, clarity, objectivity, abstraction, generalization..... Our preference is to continue with synthesizing, fixedness, and homologation. We do not want to focus on cunning and instead try to define emotionality scientifically. In addition, for lexical and syntactic characteristics, the metamorphoses, according to Meeuwisse *et al.* (2011), should be consulted: .... Therefore, you get the double advantage of clarity by using the specific term and by reducing the ambiguities that arise, for example, from the use of pronouns, allowing you to be identified as belonging to the scientific or professional group that shares the language. Regarding the structure of these languages, it is observed that, from a syntactic point of view, the criteria of economy, precision and objectivity are strongly strengthened by frequent use of nouns instead of verbs and, always with respect for them, the reduction of time and modes, the preference for impersonal and passive ways to highlight events instead of agents.

Based on these premises, the research team returns to the design phase carried out by the LSP text producer and performs the following processes: identify the operator in a certain number of possible sentences, that is, identify and then submit them to their skillful use by describing the possible combinations of verbs with forms of names (experimenting with the concrete possibilities of the speaker). Structuring and de-structured tests of possible economic language sentences begin, and as a result, their calculating and substitutive effects are observed, such as: Carlo lets know Dino;  $N_0$  Vsupp.  $N_1$ ; Carlo is the informant of Dino  $N_0$  Vsupp. Det.  $N_2$ ; Carlo gives information to Dino,  $N_1$  Vsupp det  $N_2$ . Certainly, it is only the transfer of a sample model of many marked sentences. In addition, a high rate of pre and post verbal pronominal effects such as: I am writing to inform you about Ppv Ppt V det V Ppt .I have the honor to inform you Ppt  $N_0$  det. V Ppo Basic phrase: Joe writes to Max  $N_0$  V det  $N_1$  = I write a sentence to you = change and dissociative use of the initial sentence = I still write to you = V Ppo, etc. In this phase of analysis, we can already define the validation of the criteria and characteristics of the ESP languages on the LG model because the minimum unit of grammar and semantics starting point of investigation on the LG studies is the Fs e From Gotti (1991). The structure of these languages meets the criteria of economy, precision, and objectivity from the syntactic point of view.

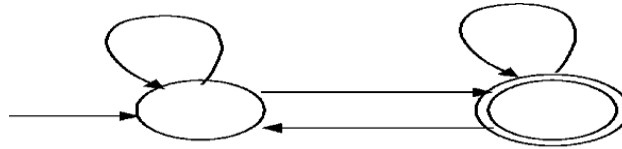
It should be noted that the Fs is a grammatical and semantic context formed by a single predicative element (Operator-modifier -Name). Thus, the verbal operator and the use of superior lexical units are essential elements, which prevail for the construction of fixed sentences and meanings and because in the LSP lexicon fixed distribution combinations are preferred for the minimum or 0 of variability and co-occurrence between words and therefore with semantic atonicity.

A smart language with the following characteristics has already been scientifically defined by the research team, with homologation, fixity, non-compositionality, synthesizing, and emotional predictability.

### **3. Manufacturer and generator acronyms: phase B**

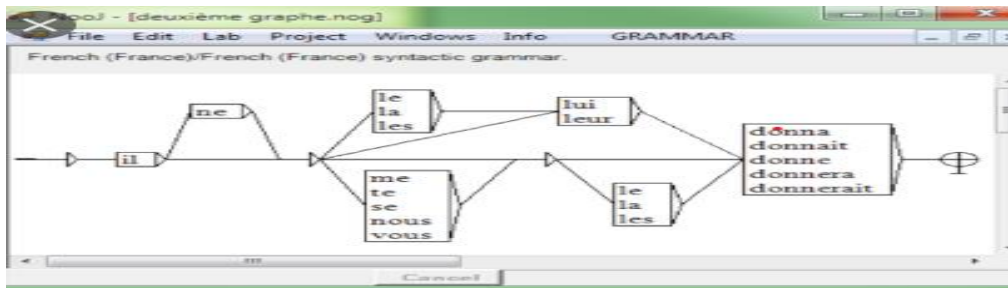
The phase of data change is represented by the acronyms manufacturer and generator, and we employ computer environments that perform scientific calculations and supercomputing to aid in designing actions. During the elaboration process for converting LSPs into LNG, humans become formal constructors, i.e. they rewrite the text in formal terms. The language being reformulated results in the fixed sentence becoming a keyword and formal element of semantic unity. The automaton's role involves describing and defining itself as a language recognition or acceptor. As from Bertoni *et al.* 2003, a finite state automaton A is a system  $A = \langle Q, \Sigma, \delta, q_0, F \rangle$ , where Q is a finite set of states,  $\Sigma$  is a finite input alphabet,  $\delta: Q \times \Sigma \rightarrow Q$  is the transition function,  $q_0 \in Q$

is the initial state, and  $F \subseteq Q$  is the set of final states. The function  $\delta : Q \times \Sigma \rightarrow Q$  can be extended to  $\delta : Q \times \Sigma^* \rightarrow Q$  as follows:  $\varepsilon \in \Sigma^*$  being the empty word, we let  $\delta(q, \varepsilon) = q$  and  $\delta(q, wa) = \delta(\delta(q, w), a)$  for any word  $w \in \Sigma^*$  and symbol  $a \in \Sigma$ . Thus, the language  $L(A)$  recognized by the automaton  $A$  is the set  $L(A) = \{ w \in \Sigma^* \mid \delta(q_0, w) \in F \}$ .



**Figure 1.** A state diagram of a finite state automaton

A new grammar, new techniques, and a different symbology have been developed. But especially NLG coding techniques related to elementary design calculations and complex calculations required for instrumentation are being developed, such as environments and products where sophisticated designs are made using NLG languages. The analysis of text in the 'generation' mode can be done by using large dictionaries and grammars with a priori structure, to generate automatic paraphrases of phrases that are described by grammars, but not written texts. NooJ is a software that helps with language development and corpus processing. NooJ is the evolution of Intex (Silberztein 1993, 1999), which Silberztein worked on in the decade 1992-2002 under the guidance of Maurice Gross at LADL. NooJ allows the construction of dictionaries and electronic grammars and their application to large corpora. Thanks to these tools it is possible to build main linguistic resources: dictionaries and grammars. The language modules concerning economic languages have been performed at the Computer Pole of the University of Salerno dir. and is the evolution of the lingware developed for Intex. NOOJ, which is the natural implementation of INTEX. It is described by Max Silberztein in Bucciarelli *et al.* (2017).



**Figure 2.** (Implementation of type 3 grammars Published in COLI

#### 4. Web crawling and software crawler per data center

Web crawling is an activity that automatically and systematically explores the Web to search for content / documents to download. From the main html page of the interface, a user can create a web crawling session, log in and monitor the generated Internet traffic. A crawler (also called a web crawler, spider, or robot) is software that analyzes the content of a network (or a database) in a methodical and automated way, usually on behalf of a search engine. Specifically, a tracker is a type of bot (program or script that automates operations), which acquires a textual copy of all documents on one or more web pages, creating an index that allows, subsequently, searching and visualization. Extremely common use is made of crawlers on the Web; it is based on a list of URLs to visit provided by the search engine (which, in turn, is initially based on the addresses suggested by the users or on a list preloaded by the programmers themselves). During the analysis of a URL, it identifies all hyperlinks present in the document and adds them to the list of URLs to visit. The process can be concluded manually or after a certain number of links have been followed. For a database and data center production, we propose a safe index based on the Bloom Filter Count (CBF) to search for multiple classified keywords. The proposed scheme uses multiple

algorithms to maintain and search for CBF, while a pruning algorithm is utilized to eliminate repeated elements and save space. In addition, the relevant scores are encrypted by the Paillier cryptosystem. The encryption of the relevant scores in different bits guarantees that they can withstand statistical analysis of the encrypted text. Likewise, since the Paillier cryptosystem supports the homomorphism addition of the encrypted text without knowledge of the private key, the main processing work in the classification could be moved from the user side to the cloud server side. Hence, the proposed scheme has enormous potential in mobile devices with limited resources, such as mobile terminals of Yao *et al.* (2016). Secure index for resource-constraint mobile devices in cloud computing. *IEEE Access*, 4, 9119-9128... In this section, we present manifold algorithms for CBF. In our scheme, we will build and lookup the SICBF index by these algorithms:

There is an array C of m counters where each counter  $C_i$  is associated with  $i$  bucket of the hash table. We compute  $k$  hash functions  $h_1(sc), h_2(sc), \dots, h_K(sc)$  on each input item and increment corresponding  $k$  counters indexed by their hash values. The item  $T$  is stored in the list  $A_i$  associated with  $k$  buckets. As shown in

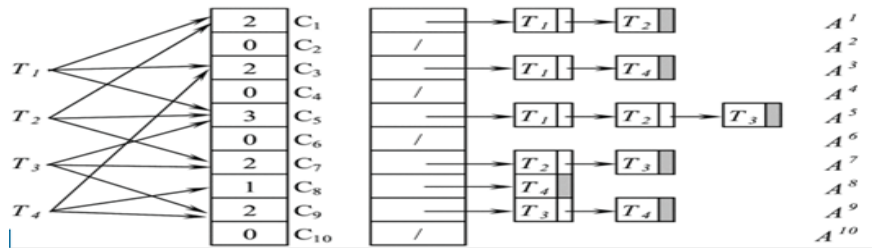


Fig 3. (Four items are inserted into CBF)

**Algorithm 1 Insert an item into CBF**

- 1: **CBF Insert Item** ( $x$ )
- 2: for  $i = 1$  to  $k$  do
- 3:  $C_{h_i(sc)} ++$
- 4:  $A^{h_i(sc)} = A^{h_i(sc)} \cup x$

Algorithm 2- Delete an item from CBF

- 1: **CBF Delete Item**( $x$ )
- 2: for  $i = 1$  to  $k$  do
- 3:  $C_{h_i(x)} --$
- 4:  $A^{h_i(x)} = A^{h_i(x)} - x$

**Algorithm 3 Search an item from CBF**

- 1: **CBF Search Item**( $x$ )
- 2: **if** ( $\cup_{i=1}^k C_{h_i(x)} > 0$ ) **then**
- 3:  $I = \text{index of min } \{ \cup_{i=1}^k C_{h_i(x)} \}$
- 4: **if** ( $x \in A^I$ ) **then**
- 5: **return**  $A^I(x)$

Other authors may disagree with the hypothesis that the team has examined and described through a descriptive process. As in the technological society, the global citizen manages to transfer their anthropologies and cultural values even through synthetic communication, it seems to contradict the characteristics that have been validated so far, but it is also said that a good manipulation of the phrase can lead to success.

**5. The text advertising the sale of emotions**

Creating a semantic analysis of the automatic process of a text advertisement is not an easy task, according to Ingale *et al.* (2012), and by presenting a specific example of the risks of producing

sentences that are not useful for text construction. Textual analysis advancements have led to the recent interest in emotion detection in computational linguistics. The use of emotion detection in textual analysis is the newest, so it has weaker standard methods. In this way, we will deal with analysis or pattern matching technique, and syntactic driven parsing technique as regards Natural language processing (NLP). NLP is the computerized approach to analyze text that is based on both a set of theories and a set of technologies. It is concerned with the interactions between computers and human (natural) languages. NLP is presenting naturally occurring texts at one or more level of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications. Like this manner, as Wiebe *et al.* (2004) point out, NLP applications such as information extraction, question answering, text categorization, and summarization could benefit from having knowledge of subjective language. Their method identifies potentially subjective collocations automatically, including fixed collocations and collocations with placeholders for unique words.



Fig 4. Syntactic driven parsing technique (Wiebe, Wilson, Bruce, Bell, and Martin, 2004)

The advertising message comes predefined with a flash of alliterative, dysmorphic, and phonic morphemes, as well as the fragile use of pronominal position, and images accompanied by sounds that convey emotions and that lead us to a fantastic world. Although it may seem utopian, it is true.

Humans calculate everything without realizing it, which creates an emotional resonance that can be explained. This is how the team is described.: the model presented is refutable according to Ingale and Kasat (2012), because the calculation of emotion production is structured to produce at the reception the same emotions that come from the words or sounds of a pure heart. Predefined advertising messages arrive between a flash of alliterative dysmorphic morphologies, a fragile position exists in the use of pronouns because they are the predefined results of the techniques of writing fixed texts from yesterday and today. For the research work, the team has divided the calculation of production: calculation techniques and laboratory techniques. In the past, a man with an elementary calculation prepared the text in accordance with the defined model. Our approach to the emotional text involved a rhetorical calculation style and the use of phrase manipulation and word juxtaposition to create tone and sound, that is, syntactic devices and morphological changes were arranged according to the order of the sentences. In classical literature, the final position was usually reserved for verbs preceded by their complements, while the subject was usually given preference at the beginning (VS). This is utilized to highlight certain parts of a sentence or to produce a sequence of words that meets the needs of euphony or expressiveness. Latin has already taught us that word order is not unimportant: on one hand, it is controlled by specific habits and preferences; on the other, it is governed by considerations of meaning, style, or rhythm. The use of the verb in the initial position in a war chronicle is a consequence of this functional interpretation, the *Bellum Africorn* where it appears that the verb at the start of the sentence represents a clear approach strategy. This position arises when the sentence is not self-sufficient, but rather serves as a close statement, resulting in strict concatenations.

*Accidit res incredibilis, ut equites minus XXX Galli Maurorum equitum. 11 Hispania loco pellerent fugarentque in oppidum (6.3).*

We use the same production techniques nowadays, but with instruments that are increasingly perfect for converting words, sounds, movements, and emotions. If calculated well, sound, and morphemic effects can be produced even in fixed texts when the speaker's phonology is considered, which involves an exhaustive analysis of words and sounds. This is to validate the transformation of LSP languages into ESP. The study ends with a laboratory analysis of cause and effect. The project focuses on the psychoeducation of the soul, ranging from perception analysis to automatic elaboration calculations. The team's demonstration shows that decoding is done using various techniques established by linguistic and professional codes, the emotional degree and tone levels of the communicative act are personalized by the mind. In fact, it is in the structuring of the text that the degree, level, and methods of structuring the text are calculated first. In this phase, it defines a precise emotion as love, and searches for words, sentences and uses and elaborates them, that is, a predefined message is projected, such as sheet music, writing, sound, signs, traction, brush, and so on. If we describe a tonal, phonic degree, the density of color and pressure, the deictic position, dictated by the soul at the time of perception, will make you rejoice and suffer, since this belongs only to man, to his sacred values, to his deontology and its values. Despite being predisposed to emotional messages reaching his/her heart, a man/woman can produce the same messages or likely matching the emotions dictated by the heart. The experiment will demonstrate that even calculated movements can elicit emotional responses. These statements claim that the soul's voice can be heard even by those who are deaf.

Let  $C$  be an elastic string of length  $l$ , density  $\mu$ , tensioned by two opposing forces, applied to its ends, of intensity  $T$ . Think, for example, of the string of a guitar.

If we pinch  $C$ , that is, we move one of its points  $P$  from the equilibrium position, due to the hypothesis of elasticity of  $C$  and the surrounding medium, the string oscillates and generates a sound (musical note) whose fundamental frequency is given by:

$$= \frac{1}{2l} \sqrt{\frac{T}{\mu}}$$

This relationship establishes a link between the frequency emitted by the string and the quantities (length, density, voltage) that characterize it, so that by varying at least one of these parameters we obtain oscillations with different frequencies and, therefore, different sounds (or notes: LA  $f = 440$  Hz).

Let now us consider two chains  $C1$  and  $C2$  that have the same length, density, and tension.

Let us suppose we only vibrate the first string  $C1$ . After a short period of time, the second  $C2$  string will also begin to oscillate with the same frequency, emitting the same sound (or note) as  $C1$ . To verify it, just stop the oscillation of  $C1$  and you will continue to hear the same sound, this time coming only from  $C2$ .

Hence, the phenomenon is given the name of resonance and the relative frequency, that of: Resonance Frequency. In doing so, we could say that the condition of existence of this phenomenon is that the two bodies have the same resonance frequency. Man/woman manages to create these conditions. This is the magic of a virtual world.

## **6. Lexical-algebraic grammar in the NooJ environment**

Having to proceed with validation in a linguistic environment or formal coding of a new formal grammar and dictionary, and implement co-occurrence texts, we will move on to the realization of transformational and distributional analysis of the sentences tested in the chosen traits, with a local grammar and an electronic dictionary in a NooJ linguistic environment. Silberztein's work on Intex (Silberztein 1993, 1999) under Maurice Gross at LADL resulted in the evolution of NooJ over a decade (1992-2002). Various types of linguistic descriptions can be formalized by linguists using NooJ: spelling dictionaries, linguistic dictionaries that focus on simple words, multi-word units and frozen expressions, flexible and derivational morphology, local, structural, and transformational syntax. An important feature of NooJ is that all linguistic descriptions are reversible, i.e. both a parser (to recognize sentences) and a generator (to produce sentences) can



use them. Like this manner, in line with (Silberztein 2012, 2016), we can show and build, by combining a parser and a generator and applying them to a syntactic grammar, a system that takes a sentence as input, and which produces all the sentences that share the sentence the same lexical material with the original expression.

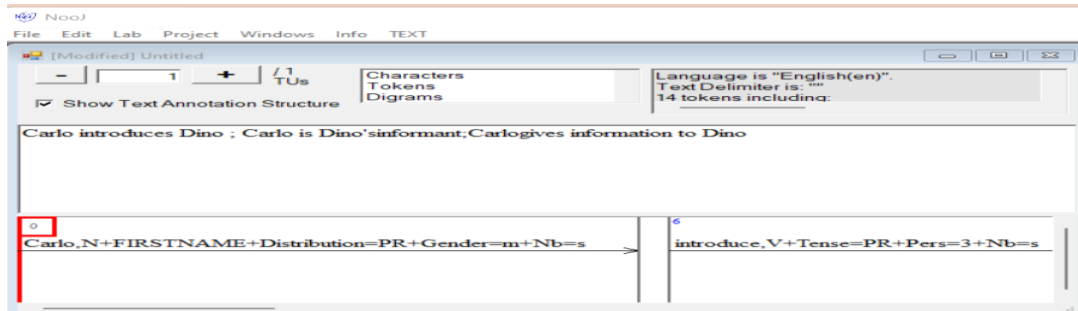


Figure 5. Minimum sentence Fs

The second one can be implemented in NooJ via the following grammar:

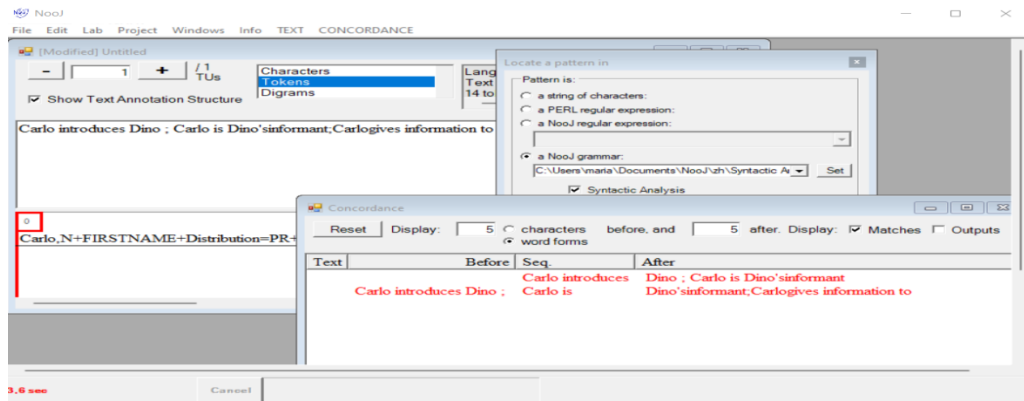


Figure 6. Sentence processed LOC

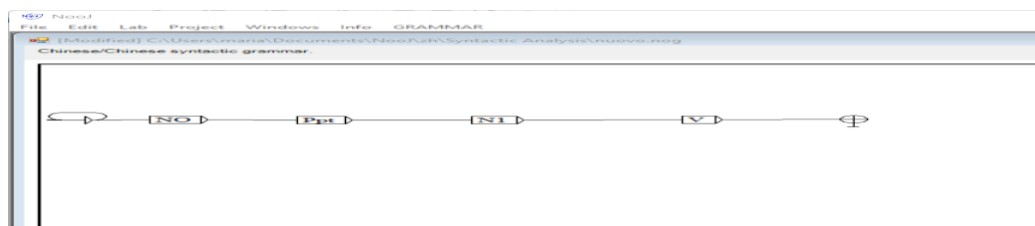
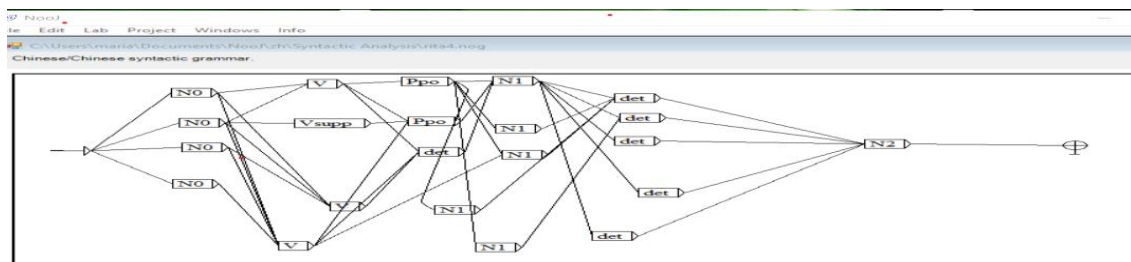


Figure 7. Implemented Sentence LOC



**Figure 8.** Variables of sentence analysis

This chart uses three variables \$ NO, \$ V, and \$ N1. When analyzing the sentence, *Mercy lives in you*, the variable \$ NO stores the word, *Mercy*. \$ V stores living \$ N1 stores in you. The output of the grammar "\$ N1 is \$ V\_V + PP of \$ NO" produces the string *Mercy lives in you*. Note that morphological operations such as "\$ V\_V + PP" operate on NooJ's Atomic Linguistic Units (ALU) rather than just strings. In other words, NooJ knows that the word form is an instance of the verb to live. It can produce all the word form conjugated and derived from this ALU, such as the living. Here, \$ V\_V + PP takes the value of the variable \$ V (*ama*), lemmatizes it (*vive*), produces all its verb forms, and selects those that have properties + PP. The transformational and distributional analyzes of the minimum sentence of the chosen triplet has a fixed formal code and represents the minimum unit that...

## **7. The digital text workshop in the NooJ environment**

The lexical-grammatical and automatic text analysis: <sup>1</sup>

*NooJ is geared towards developing grammars processing large amounts of texts and therefore contains a full-fledged corpus processing module, indexing, annotation and querying of corpora is a basic functionality. Corpora's size is typically .....*

*Text can be imported in a wide variety of formats and a lexical analysis is immediately applied based on a robust dictionary module that has a built-in morphological analyzer. The result of the lexical analysis becomes the initial tier in a set of stand-off annotation levels containing at first POS and morphological codes as well as the result of any morphological grammars that carried out typical pre-processing normalizations of the text. This basic annotation, amounting to indexing of the corpus, can be enhanced with an arbitrary number of further tiers as subsequent syntactic grammars are applied in a cascaded manner.....<sup>2</sup>*

Transformational analysis of feature mechanisms and graphs are used to produce computations and graphs in the NooJ environment:

*One of the most attractive features of the system that immediately appeals to users is the ease with which sophisticated grammars (of various levels of computing power) can be built in graph form and applied to corpora as a query (Beesley & Karttunen, 2003) (see, Figure 2, for an example). This feature alone makes NooJ ideal for teaching and rapid application development tool alike. The graphs have alternative textual notation in case that mode of definition proves more applicable<sup>3</sup>*

---

<sup>1</sup> Per la realizzazione dell'analisi testuale automatica è stato utilizzato il pacchetto software NOOJ, principalmente basato sull'uso di dizionari elettronici, grammatiche locali e automi e trasduttori a stati finiti (da qui in poi FST). Nel corso delle applicazioni del software, sono stati utilizzati i dizionari elettronici sviluppati dal Dipartimento di Scienze della Comunicazione dell'Università di Salerno, ovvero cinque dizionari del sistema DELA (dizionari elettronici di parole semplici e composte) ai quali è stato affiancato un dizionario elettronico specialistico del settore della comunicazione .

<sup>2</sup>Max SILBERZTEIN1 Tamás VÁRADI<sup>2</sup> Marko TADIĆ<sup>3</sup> Open source multi-platform NooJ for NLP

<sup>3</sup> IBIDEN pag

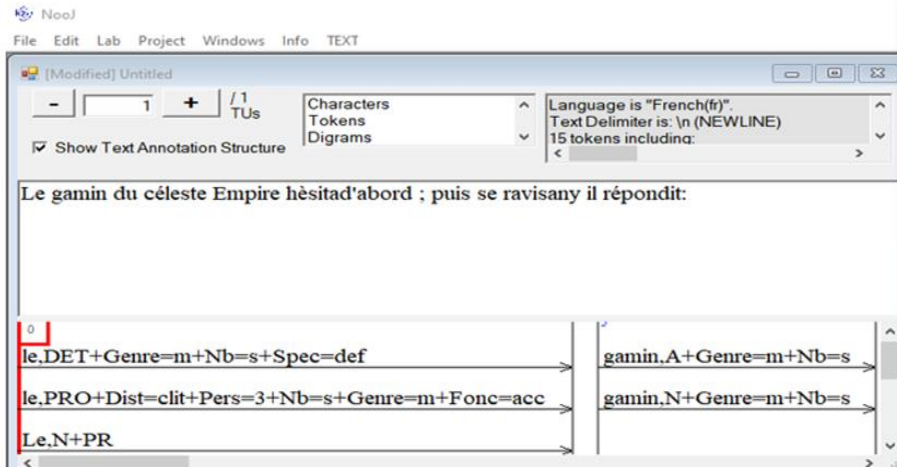


Figure 9. Transformational analysis: Pronominalization

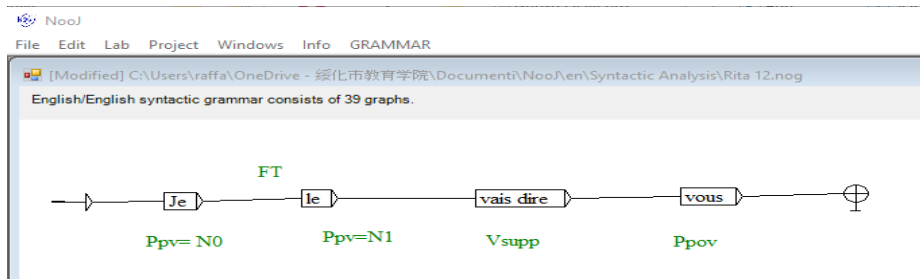


Figure 10. Fixed Phrases

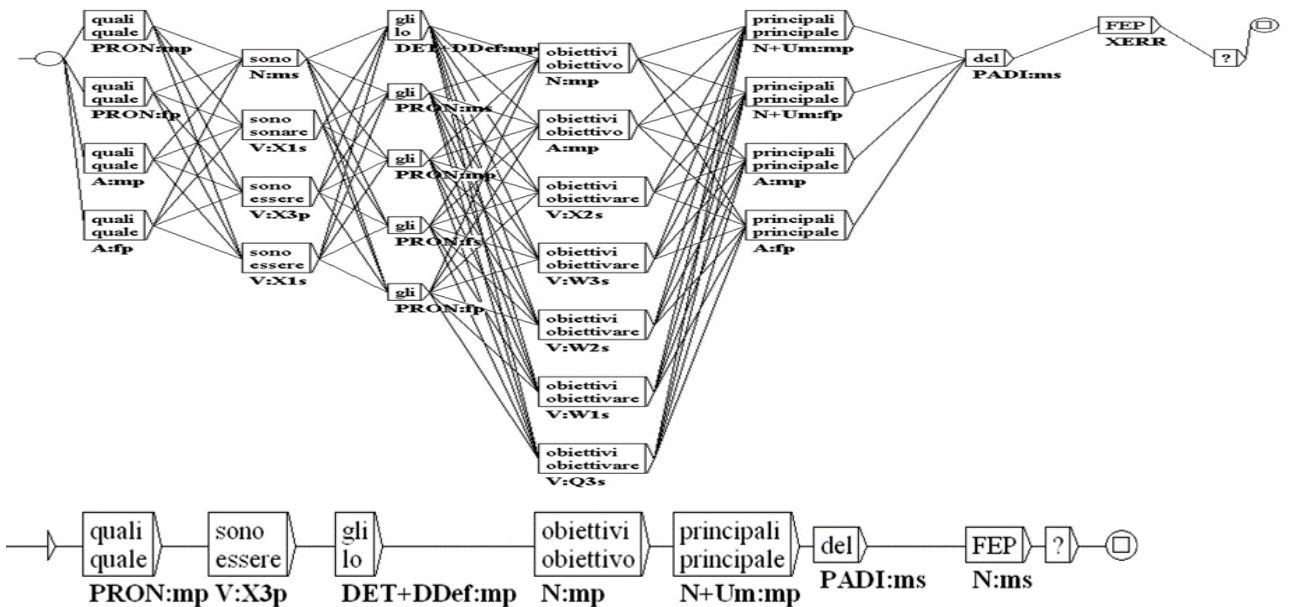


Figure 11. Results of Pronominalization and Fixed Phrases

In accordance with the above, Silberztein (2005) points out that one characteristic of NooJ is that its corpus processing engine uses large-coverage linguistic lexical and syntactic resources, which allows NooJ users to perform sophisticated queries, which include any of the available

morphological, lexical, or syntactic properties. In this way, these Silberztein methods could serve as a basis for carrying out operations in the Linguistic, Physics, and Mathematics setting that would allow users to build broader mathematical, lexical, and semantic structures in these fields of study.

As a great complement to these approaches, we could also make use of *PeCo-Graph*<sup>4</sup>, which is the software for obtaining and graphically representing the counting data obtained by the Visual Tools counting applications. It does draw on a simple PC application that, on the one hand, facilitates data collection and, on the other hand, simplifies its analysis and graphical representations. In other words, it is a multilingual application for computer databases.

## 8. Conclusion

The results of the research have led us to validate the moments of the design, because it is in that phase that the mind builds the semantic contents and prepares with methods, tools, and techniques the purpose, criteria, and tools of the communicative message. In this phase, the man/woman has innate competences that make it inevitable for them to transmit smells and perfumes, his/her land has history, anthropology, and culture, along with his/her ego, which is about his/her identity and acquired competences. In the first phase of design man/woman uses linguistic tools of his thought and manipulates language, using only his/her skills and identity, but above all prepares the text for use and purpose. He/she then moves on to processing and generates a finished product using complex databases and calculations. With proper design, the reaction will be cause-and-effect, just like written text, and the risks are identical. A professional emotional advertising text, if well produced, gives the same emotions, and invites us to buy the product. For this reason, Silvia Bernardini's thesis is irrefutable when she states that today the global citizen sends messages in ESP that is tonal and phonic in perfect technicalities. It is the grammar of feelings that prevails and guides the technique, but that makes us rejoice, love, laugh and cry, because it is only the man/woman, who can...

## 9. Bibliography

- Beesley, K. R., & Karttunen, L. (2003). Finite-state morphology: Xerox tools and techniques. *CSLI, Stanford*, 359-375.
- Bernardini, S. (2004). Corpora in the classroom. *How to use corpora in language teaching*, 12, 15-36.
- Bertoni, A., Mereghetti, C., & Palano, B. (2003, July). Quantum computing: 1-way quantum automata. In *International Conference on Developments in Language Theory* (pp. 1-20). Springer, Berlin, Heidelberg.
- Bucciarelli, R., & Elia, A. (2006). *La grammatica dell'italiano: comparazioni e descrizioni dei codici linguistici*.
- Bucciarelli, R., Capone, R., Tortoriello, F. S., Greco, M., Savarese, G., & Enriquez, J. (2019). Learning Analytics-Scientific Description and Heuristic Validation of Languages NLG. *Journal of e-Learning and Knowledge Society*, 15(3), 251-261.
- Bucciarelli, R., Falco, V., & Savarese, G. (2018). Risorse linguistiche per la generazione automatica della lingua dei segni naturali. *FORMAZIONE & INSEGNAMENTO. Rivista internazionale di Scienze dell'educazione e della formazione*, 16(2), 13-24.
- Bucciarelli, R., Falco, V., & Savarese, G. Risorse linguistiche per la generazione automatica della lingua dei segni naturali *Linguistic Resources for Automatic Natural Sign Language Generation. La formazione dei talenti Prospettive di ricerca*, 15.
- Bucciarelli, R., Marcone, R., Santoro, F., & Dolci, C. (2017). *I Technology and languages*

---

<sup>4</sup> <http://web2.visual-tools.com/downloads/PeCoGraph-UserManual-en.pdf>

- of the new economy. Salerno IRIS In-Text (Bucciarelli, 2017).
- Bucciarelli, R., Elia, A., Landi, A. (2000) *Dalla Grammatica al testo poetico: una lezione di linguistica*, Napoli, Liguori.
- Elia, A., Martinelli, M., D'Agostino, E. (1981), *Lessico e strutture sintattiche*, Napoli, Liguori (EMDA)
- Gotti, M. (1991). *I linguaggi specialistici: caratteristiche linguistiche e criteri pragmatici*. La Nuova Italia.
- Gross, M. (1999). A bootstrap method for constructing local grammars. In *Proceedings of the Symposium on Contemporary Mathematics* (pp. 229-250). University of Belgrad.
- Hoffman, R. R. (1984). Recent psycholinguistic research on figurative language. *Annals of the New York Academy of Sciences*, 433, 137–166.
- Ingale, H. T., & Kasat, N. N. (2012). Automated Irrigation System. *International Journal of Engineering Research and Development*, 4(11), 51-54.
- International Conference on High Performance Computing & Simulation (HPCS) (pp. 1030-1033). IEEE.
- Meeuwisse, A., Mariano, A., D'Amato, G., Ambrosino, F., Aprea, G., Scaramuzzino, R., & Swärd, H. (2011). Everyday realities and visionary ideals among social workers in the Nordic countries: A matter of specialization and work tasks? *Nordic Social Work Research*, 1(1), 5-23.
- Silberztein, M. (1993). Dictionnaires électroniques et analyse automatique de textes: le système INTEX. (*No Title*).
- Silberztein, M. (1999). Text indexation with INTEX. *Computers and the Humanities*, 33(3), 265-280.
- Silberztein, M. (2005). "NooJ: a linguistic annotation system for corpus processing." *Proceedings of HLT/EMNLP 2005 Interactive Demonstrations*.
- Silberztein, Max. (2012). "Variable unification in NooJ V3." *Automatic Processing of Various Levels of Linguistic Phenomena: Selected Papers from the NooJ 2011 International Conference*. Cambridge Scholars Publishing.
- Silberztein, M. (2016). *Formalizing natural languages: The NooJ approach*. John Wiley & Sons.
- Wiebe, J., Wilson, T., Bruce, R., Bell, M., & Martin, M. (2004). Learning subjective language. *Computational linguistics*, 30(3), 277-308.
- Yao, H., Xing, N., Zhou, J., & Xia, Z. (2016). Secure index for resource-constraint mobile devices in cloud computing. *IEEE Access*, 4, 9119-9128.

EDOC@Work 3.0. <https://www.edocwork.it/>  
<https://rscaramuzzino.wordpress.com/2016/08/20/i-linguaggi-specialistici/>  
 EDOC@Work 3.0. <https://www.edocwork.it/>.