Semantic Predicates in the Business Language

Maddalena della Volpe[1(✉)], Annibale Elia[2], and Francesca Esposito[2]

[1] Department of Business, Management and Innovation System, University of Salerno, Fisciano, Italy
mdellavolpe@unisa.it
[2] Department of Political, Social and Communication Sciences, University of Salerno, Fisciano, Italy
{elia, fraesposito}@unisa.it

Abstract. In recent years, the interest in the use of language for business has grown. It is recognized that the hidden persuasive linguistic potential improves the company's positioning in the public consciousness. The language of the business world is multifarious: we try to identify its features and behaviour, considering the evolution that it has faced primarily with the globalization of markets. Business activities are so complex that they require the application of several disciplines at the same time and therefore the use of specific languages and technical terminology. In order to reach an efficient analysis of business language, this study explores the role of semantic predicates constructed from lexical and the syntactic structures in which they are placed within business communication contexts. From the point of view of LG framework, a set of lexical-syntactic structures defines the value of semantic predicates, while the arguments selected by each semantic predicate are given the value of actants, subjects included. The features of each verb are expressed by the application of the rules of co-occurrence and selection restriction, through which verbs select semantically their arguments to construct acceptable simple sentences. In this way, the entries belonging to electronic dictionaries should be classified presuming their similarity and proximity. Even if the list of semantic tags is not simply identifiable, grammars could be built for single sets of semantic predicates. LG descriptions assign correlated predicates and arguments by applying electronic dictionaries of Italian. Using NooJ environment and Italian linguistic resources to automatically processing natural language, we will process a corpus of business documents. We will show and describe the syntactic structures, semantic and syntactic properties of predicates, in order to build formal grammar for business language.

Keywords: Semantic predicates · Business language
Natural language processing · Text Mining · NooJ application

1 Introduction

In the business language, many special expressions are used to define and describe the actions of a company within itself or with the outside world. There is often only one verb to express a context, a process or an action. In this article, we will analyse the most common predicates in business documents to understand their functions and features.

S. Mbarki et al. (Eds.): NooJ 2017, CCIS 811, pp. 108–116, 2018.
https://doi.org/10.1007/978-3-319-73420-0_9

Moreover, through an automatic linguistic analysis, it is possible to verify the influence of the co-occurrences of these predicates in order to understand the text. By adopting the bipartition between operators and arguments, first proposed by Harris [1, 2], and subsequently adopted by Gross [3–9], we can assert that predicates assume the ability to request specific arguments and constitute a potential simple sentence. Although not verb-centric, Lexicon-Grammar, containing all the possible combinations of simple sentences at the distribution level, considers semantic information that allows us to recognize the predicates and their topics (operators) at a semantic level [10]. From the point of view of LG framework, a set of lexical-syntactic structures defines the value of semantic predicates, while the arguments selected by each semantic predicate are given the value of actants, subjects included [11]. The features of each verb are expressed by the application of the rules of co-occurrence and selection restriction, through which verbs select semantically their arguments to construct acceptable simple sentences. According to Monteleone and Vietri [12], we have semantic predicates expressing the intuitive notion of "exchange" (Transfer Predicates), "motion" (Movement Predicates) or production (Creation Predicates). Each set of semantic predicates assumes those arguments with which they have compatible semantic roles.

Transfer Predicates have a "giver", an "object to transfer" and a "receiver", as in the sentences:

1. Mario (giver) gave a cake (object to transfer) to Juliet (receiver)
2. Juliet (receiver) received a cake (object to transfer) from Mario (giver).

Movement Predicates select an "agent of motion", "object to move" and a "locative name", as in the following:

1. Mario (agent of motion) went to Paris (locative name)
2. Mario (agent of motion) moved the books (object to move) from his house to the office (locative names).

Creation Predicates, finally assume a "creator" and a "creation":

1. Mario (creator) wrote a novel (creation)
2. Juliet (creator) composed a song (creation).

In this way, the entries belonging to electronic dictionaries should be classified presuming their similarity and proximity to semantic predicates. Even if the list of semantic tags is not simply identifiable, due to the polysemy of simple nouns, grammars could be also built for single sets of semantic predicates. LG descriptions assign correlated predicates and arguments by applying electronic dictionaries of Italian. It is also possible to build grammars that annotate all specific semantic predicates. In the following paragraphs, we analyse the language of business, identifying features and singularity. Business documents define the complex world of enterprise and describe business activities, functions and actors. We analyse business plans to explore the use of language, and in particular in this study we focus on semantic predicates. According to Elia [10], we take into consideration LEG-Semantic Role Labelling system (LEG-SRL) for Italian, built on 2000 verbal uses, included in semantic predicates classes. We recognize some verbal uses recurring in the documents, with the purpose of explaining and improving the companies express themselves relying on communication exercise.

2 The Language of Business

As far as the business language is concerned, we must consider two fundamental aspects that make the analysis rather complex. To express business activities in their complexity, as well as in their diversity, we have to consider, on one hand, the sub-languages that characterize this world, and on the other hand terminology. For instance, sublanguages are used to describe professional activities belonging to different business sectors: banking, trading, accounting, communication, logistics, administration etc. Another issue is referred as terminology: no one could say that business has a specific and limited vocabulary. The study of language in business contexts is highly inter-disciplinary [13]. Business activities are so complex that they require the application of several disciplines at the same time, and therefore the use of specific languages. Although, it is always necessary that the circumstances, in which terms are uttered, should be in some way, or ways, appropriate. The combination of business functions and processes is impacted by improved communication. From company to company, we have seen language skills consistently deliver tangible business value and virtuous results for organizations that invest in language training.

Ford and Wang [14] observed how the use of language in the field of strategic management has been the subject of many studies [15–18] just because there is no unique classification of words as it exists for other disciplines such as Economy. Every strategic document is a stream of decisions [19] and actions whereby it does not just describe reality but performs it in the same moment in which they are representing it.

The language of the business world is definitely multifarious: we have tried to identify its features and behaviour, considering the evolution that it has faced primarily with the globalization of markets. In the last thirty years, the interest of researchers in the variety of specific language uses has increased significantly [20]. However, in relation to different specialized varieties of the language, there is no a unified terminology, and tags used in this field of research by various researchers are different. Nevertheless, we must consider the fact that the use of certain terms entered in the common language through mass media, as we know, often becomes the point of contact between the specialists and the people. Thus, we will have a kind of coded language that is typically used in the field of the economy, and another type of language that instead has developed among the experts, a type of jargon, which then became part of everyday life through the media.

For instance, in previous studies [21, 22] we dealt with the specific lexicon used by media to describe the phenomenon of startup companies. We studied how the Italian terminology and this specialty language can be used in routine automatic text analysis. Using NooJ environment [23–28] for the automatic processing of natural language, through the application of electronic dictionaries of terminology and specialty, we analysed a corpus of 2000 journal texts centred on the startups topics. After the analysis, we collected about 400 entries, a great part of which belongs to the semantic fields of economics and informatics and a small part to professionals, revenue and law. Moreover, it appears that the terminology of the world of startups is rich of foreign words, coming mainly from the United States. Through the study of the presence, frequency and origin of lexical entries, it is possible to grasp certain phenomena implicitly expressed in the texts analyzed, with the objective of a better understanding

of the evolution of the ecosystem of startups. The specialty language that has been determined requires the continuous, online monitoring of a dynamic and innovative vision in the specialized terminology field. On the other hand, it derives from the fact that there is a very strong presence of borrowings in the English language in the lexicon of startups. This data could be taken as an invitation to extend the research by adapting these terms to the Italian language system, thus satisfying the need to find effective correspondents to describe certain concepts. This case shows even how mostly technical words enter on our common language through mass media, and become our opportunity to comment some socio-economic events.

Nowadays, in the language of business, we can identify two level of language: specialized and popular. The specialty language includes all the features of the sector language, while the popular language is spread through mass media. The popular level resorted to some mitigations, making less complex the language, or recourse to metaphors. Predominantly, economic dictionaries characterize the language of Business, but the enterprise system is so complex that it naturally requires the intervention of more specialized languages in the interaction processes, based on the nature of the enterprise and on the market in which it operates. The recognition of economic terminology is revealed only as the basis for a larger study that may involve other types of specialized language processing, within the analysis of textual documents that provide information to support strategic decisions. Thus, the language of business is partially the language of Economics, as it uses many words that have a dramatic nuance ("crisieconomica" as "economic crisis") or military origin ("manovra finanziaria" as "budgetary manoeuvre") as shown by Parantainen [29]. The most striking feature of the business language in Italian is the presence of foreign words and expressions, especially of English origin, so abounding of technicalities and terms that are often incomprehensible to the experts. To obtain an efficient Text Mining system and to apply it to the business document analysis, we have to consider typical economic language, opening our analysis' field to other knowledge domains.

Business documents are files that provide details related to a company, in fact, they are used to communicate, transact business and analyse productivity. In the meantime, business documents provide the profile of an organization and may be referred to for years to come: it is very important that they are well prepared, to avoid conveying a negative impression about the person who wrote it or the company for which it is written. Thus, writing excellent business documents is imperative for any working professional: they can be digital, occurring as electronic files, or in a physical form, written or printed on paper. Business documents range from brief email messages to complex legal agreements. Some documents are prepared by employees and business owners, while others are drafted by professionals from outside the company, such as accountants and lawyers. The most important external and internal business documents are:

(a) Business plans
(b) Letters, mail and memorandum
(c) Business reports
(d) Financial and accounting documents
(e) Operational documents
(f) Customer documents.

After choosing documents' types that we would process, we proceed with pre-processing of unstructured linguistic data. This phase goes through the application of LG theory and methodologies formalization of language (LG tables, electronic dictionaries and local grammars). Then we can process the texts in NLP software environment. After this linguistic pre-processing phase, we obtain several results, which can be integrated into different business applications [30].

3 Semantic Predicates and Syntactic Structures Groupings

At this point, we provide an example of business document automatic analysis based on LG framework: we analyse a corpus of Business Plans, recognizing a set of semantic predicates used in business language. Subsequently, we create New Local Grammars and other tools, developing a complex system that allows understanding the features of language used by expert in this field. Corpus exploration in Fig. 1, leads us to recognize a substantial number of operators that present these arguments and that provide indications of circumstantial nature.

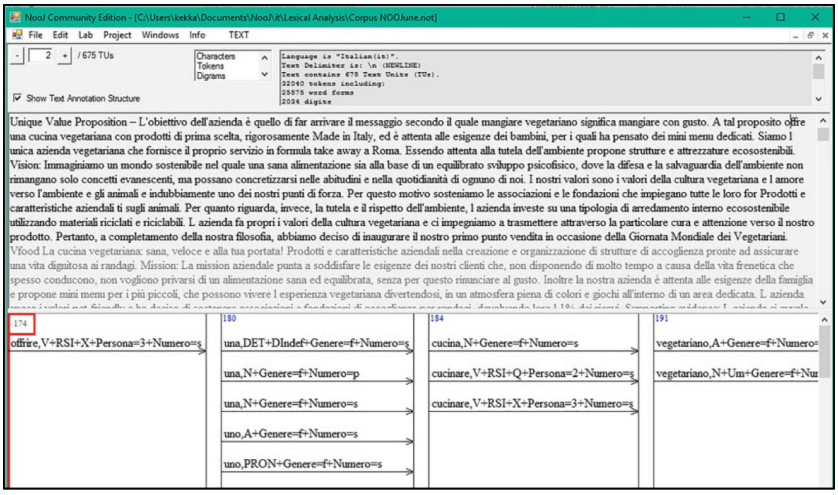

Fig. 1. Corpus exploration

The most frequent predicates in the corpus (about 100 Kb) are 21. We excluded from our observations the verbs that play only a supporting function for other verbs (to be or to have). We noticed that often some verbs are interchangeable with each other, in the sense that they have the ability to select the same lexical material that may co-occur with them, so we grouped them according to their behaviour (Table 1).

We can observe that as for the mentioned verbs, the action typically passes directly from the enterprise to the object (person, animal or thing) that receives or suffers it. Depending on their behaviour, we could associate these predicates to the grammatical classes already recognized in some previous studies [10]. Following the same examples taken from our corpus (Tables 2, 3, 4 and 5).

Table 1. Semantic predicates groupings

Transfer predicates N0 V N1 a N2	Offrire, vendere, distribuire, commerciare
Causative locative predicates N0 V N1 Loc N2	Posizionare, immettere, inserire, Introdurre
Communication predicates N0 V a N1	Proporre, presentare, garantire, assicurare
Creative predicates N0 V N1 = -um	Sviluppare, accrescere, espandere, potenziare, incrementare

Table 2. Transfer predicates (N0 V N1 a N2)

Agent giver	*L'azienda* offre strutture ecosostenibili agli ospiti
	The company offers environmentally friendly facilities to guests
Object of transfer	L'azienda offre *strutture ecosostenibili* agli ospiti
	The company offers *environmentally friendly facilities* to guests
Benef./receiver	L'azienda offre strutture ecosostenibili *agli ospiti*
	The company offers environmentally friendly facilities **to guests**

Table 3. Communication predicates (N0 V a N1)

Agent issuer	L'*azienda* garantisce la massima genuinità dei prodotti al cliente
	The company guarantees the highest genuineness of the products to the customer
Topic/message	L'azienda garantisce *la massima genuinità* dei prodotti al cliente
	The company guarantees *the highest genuineness* of the products to the customer
Benef./ receiver	L'azienda garantisce la massima genuinità dei prodotti *al cliente*
	The company guarantees the highest genuineness of the products *to the customer*

Table 4. Causative locative predicates (N0 V N1 Loc N2)

Agent	*L'azienda* immette un prodotto innovativo nel mercato
	The company introduces an innovative product into the
Place	L'azienda immette un prodotto innovativo *nel mercato*
	The company introduces an innovative product *into the market*

Table 5. Creative predicates (N0 V N1 = -um)

Agent creator	*L'azienda* accresce il fatturato
	The company increases the turnover
Topic/obj. of creation	The company increases the turnover
	L'azienda accresce *il fatturato*

The examples presented here are only classification principles, but such verbs with their uses, appear extremely frequent in the business language. We have created a local grammar on the basis of the most frequent simple sentence form in Italian [31] that we have found in the corpus. Some examples of local grammars are represented for transfer predicates in Fig. 2, communicative predicates in Fig. 3 and causative locative predicates Fig. 4. By conducting some experiments with NooJ, it is possible to label predicates and arguments to question the machine, with respect to the nature of the attendant, and of the main themes.

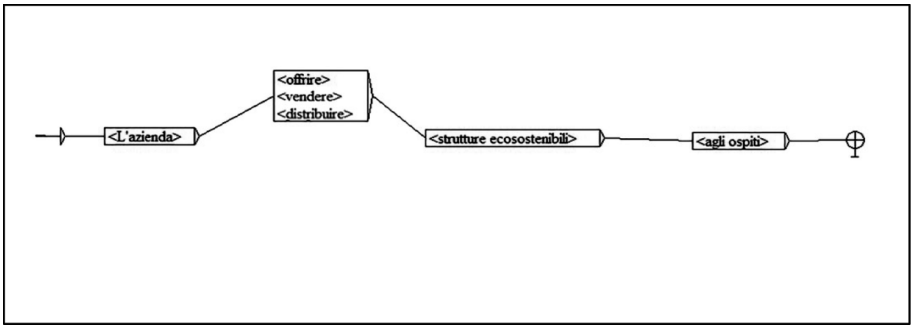

Fig. 2. Example of local grammar with transfer predicates

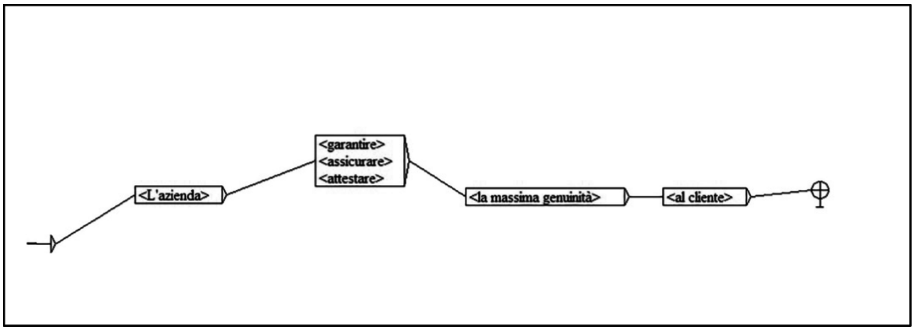

Fig. 3. Example of local grammar with communication predicates

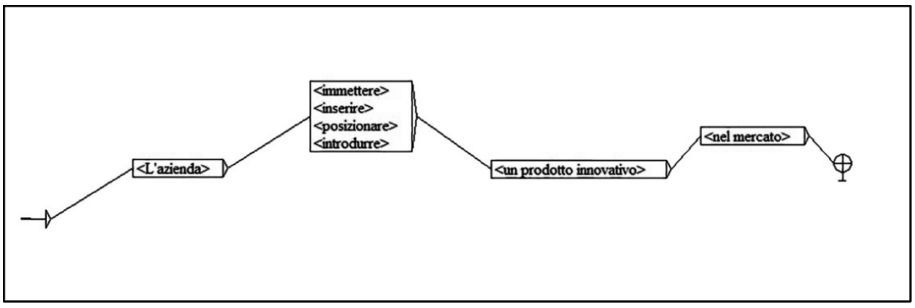

Fig. 4. Example of local grammar with causative locative predicates

4 Conclusions

To be competitive in the market and face innovation challenges, companies need to acquire specific knowledge, growing and communicating outside their values. Despite the detailed level of the methodological and theoretical framework provided, which have given us great hopes: the analysis of results has made us realise that the formalization of all linguistic phenomena is extremely complex. As we have tried to show in this paper, the exhaustive description of the lexicon and grammatical uses of a language, associated with a morphosyntactic electronic dictionary and a variety of local grammars could give satisfying results at this primary level. Semantic predicates could be used to analyse business processes, arguing that content of a text is unlabelled in advance, such as business plans, emails, and business formal communications. We admit that this study is a primary attempt to the development of a linguistic support to embed inside decision-making, with a particular reference to the document-driven analysis.

References

1. Harris, Z.S.: Distributional structure. WORD **10**, 146–162. Reprinted in Fodor, J., Katz, J.: The Structure of Language: Readings in the Philosophy of Language. Prentice-hall, Upper Saddle River (1964)
2. Harris, Z.S.: A Theory of Language and Information: A Mathematical Approach. Clarendon Press, Oxford, New York (1991)
3. Gross, M.: L'emploi des modèles en linguistique. Langages 9, pp. 3–8. Larousse, Paris (1968)
4. Gross, M.: Méthodes en syntaxe. Hermann, Paris (1975)
5. Gross, M.: Mathematical Models of Language. Prentice-Hall, Englewood Cliffs (1972)
6. Gross, M.: Méthodes en syntaxe, régime des constructions complétives. Hermann, Paris (1975)
7. Gross, M.: Les bases empiriques de la notion de prédicat sémantique. Langages, 63. Larousse, Paris (1981)
8. Gross, M.: Lexicon-Grammar. The Representation of Compound Words. In: AA. VV., Proceedings of COLING-1986, pp. 1–6. University of Bonn, Bonn (1986)
9. Gross, M.: La construction de dictionnaires électroniques, dans AA. VV. Ann. des Télécommun. **44**(1), 4–19 (1989). CNET: Issy-les-Moulineaux/Lannion
10. Elia, A.: Operatori, argomenti e il sistema "LEG-Semantic Role Labelling" dell'italiano. In: Relazioni irresistibili Pisa, pp. 105–118. ETS (2014)
11. Elia, A., Vietri, S., Monteleone, M., Marano, F.: Data mining modular software system. In: SWWS2010 – Proceedings of the 2010 International Conference on Semantic Web & Web Services, Las Vegas, Nevada, USA, 12–15 July 2010, pp. 127–133. CSREA Press (2010)
12. Vietri, S., Monteleone, M.: The NooJ english dictionary. In: Formalising Natural Languages with NooJ 2013: Selected Papers from the NooJ 2013 International Conference 12 Back Chapman Street, Newcastle upon Tyne, NE6 2XX, pp. 69–86. Cambridge Scholars Publishing (2014)
13. Studer, P.: Linguistics applied to business contexts: an interview with Patrick Studer. ReVEL **11**(21), 187–202 (2013)

14. Ford, E.W., Wang, Z.: Tackling the confusing words of strategy: effective use of key words for publication impact. Bus. Manag. Strategy, **5**(1) (2014)
15. Hoskisson, R.E., Hitt, M.A., Wan, W.P., Yiu, D.: Theory and research in strategic management: swings of a pendulum. J. Manag. **25**(3), 417–456 (1999)
16. Leontiades, M.: The confusing words of business policy. Acad. Manag. Rev. **7**(1), 45–48 (1982)
17. Nicolai, A.T., Dautwiz, J.M.: Fuzziness in action: what consequences has the linguistic ambiguity of the core competence concept for organizational usage? Br. J. Manag. **21**, 874–888 (2009). https://doi.org/10.1111/j.1467-8551.2009.00662.x
18. Ronda-Pupo, G.A., Guerras-Martin, L.Á.: Dynamics of the evolution of the strategy concept 1962–2008: a coword analysis. Strategic Manag. J. **33**, 162–188 (2012). https://doi.org/10.1002/smj.948
19. Mintzberg, H.: Patterns in strategy formation. Manag. Sci. **24**(9), 934–948 (1978)
20. Elia, A., Monteleone, M., Esposito, F.: Dictionnaires électroniques et lexique des startups. Un exemple d'analyse textuelle automatique. Dictionnaires électroniques et dictionnaires en ligne, Les Cahiers du dictionnaire **6**, 43–62 (2014)
21. Esposito, F., della Volpe, M.: Using text mining and natural language processing to support business decision: towards a NooJ application. In: Barone, L., Monteleone, M., Silberztein, M. (eds.) NooJ 2016. CCIS, vol. 667, pp. 234–245. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-55002-2_20
22. Esposito, F., Elia, A.: NooJ local grammars for innovative startup language. In: Barone, L., Monteleone, M., Silberztein, M. (eds.) NooJ 2016. CCIS, vol. 667, pp. 64–73. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-55002-2_6
23. Silberztein, M.: Nooj Manual (2003). http://www.nooj4nlp.net/NooJManual.pdf
24. Silberztein, M.: Corpus linguistics and semantic desambiguation. In: Maiello, G., Pellegrino, R. (eds.) Database, Corpora, Insegnamenti Linguistici. Linguistica n° 63, Schena Editore/Alain Baudry et C.ie, pp. 397–410 (2012)
25. Silberztein, M.: NooJ computational devices. In: Koeva, S., Mesfar, S., Silberztein, M. (eds.) Formalising Natural Languages with NooJ 2013: Selected Papers from the NooJ 2013 International Conference (Saarbrucken, Germany), pp. 01–14. Cambridge Scholars Publishing, Newcastle (2013)
26. Silberztein, M.: NooJ V4. In: Koeva, S., Mesfar, S., Silberztein, M. (eds.) Formalising Natural Languages with NooJ 2013: Selected Papers from the NooJ 2013 International Conference (Saarbrucken, Germany), pp. 01–12. Cambridge Scholars Publishing, Newcastle (2014)
27. Silberztein, M.: Analyse et generation transformationnelle avec NooJ. In: Elia, A., Iacobini, C., Voghera, M. (eds.) 2015 Proceedings of the 47th Annual Meeting of the Italian Linguistic Society "Livelli di Analisi e Fenomeni di Interfaccia". Bulzoni, Rome (2015)
28. Silberztein, M.: La formalisation des langues: l'approche de NooJ. ISTE Ed, London (2015)
29. Parantainen, P.: I prestiti non adattati nel linguaggio dell'economia, Master's degree (2001). https://jyx.jyu.fi/dspace/handle/123456789/13653
30. Esposito, F.: Semantic Technologies for Business Decision Support. Discovering Meaning with NLP Applications, Ph.D. thesis (2017)
31. Vietri, S.: Dizionari elettronici e grammatiche a stati finiti. Metodi di analisiformaledella lingua italiana. Salerno, Plectica (2008)