Technologies and new communication models for a didactic of digital intelligence ESP

Ritamaria BUCCIARELLI[1], Roberto CAPONE [2], Marianna GRECO [3], Javier Julian ENRIQUEZ [4] Francesca SANTORO [5], Raffaele MARCONE [2], Francesco Saverio TORTORIELLO [2]

[1]Università di Siena, (Si); [2]Università di Salerno (Sa); [3]MIUR, Salerno, (Sa); [4]Università di València, (Spagna); [5]Miur, Salerno, (Sa).

Abstract

Artificial intelligence (AI) is the future of mankind, which in positive terms, is compared to a "New Industrial Revolution". However, as Ross (2018) points out, the fact that an algorithm guides intimate aspects of our lives is shocking. A utopian but true idea, because in everyday life educators are "Translators" i.e. manufacturers of algorithms for teaching in the info sphere... If only we thought about search engines, which no longer give only information, but explanations, we would realize that society no longer requires skills, but rather "Excellence": The new communication belongs to the NLG languages and, therefore, the transmission is done through synthetic, standardized and non-compositional techniques and by areas...As Languages evolves, "Knowledge" is described by formal grammars and entrusted to technologies. The new communication delegates the elevation of the languages of Natural Language Generation (NLG) to the mathematical, technological and human sciences. The team follows the models: - Scaramuzzino (2016) with an explanation of the phono-morphic-syntactic and semantic characteristics of the languages LSP; Bernardini (2010) which explains the question of the replacement of L1 in L2 for the predictive detection of belonging of the ESP to the LSP. The focal parts of the research are: 1) information retrieval through the description of models; 2) retrieval techniques and new systems of text retrieval and description of digital intelligences) 3 conjecture of the epistemological status of ESP and at the same time of belonging to the LSP codes.

Keywords: *Digital intelligences, Mathematical validation*

Introduction

Every Educational Administration worldwide should be aware of the challenges and implications that the digital era might entail in the future for the Education setting, thus having to revamping it in order to meet such challenges. In doing so, Digital Era will involve a new era in the history of Linguistics or Medicine, for example. In that regard, in the context of Skills and Knowledge Acquisition for our modern global culture*, which skills are necessary for our modern global culture? How is this skill gap being solved?*

In doing so, the contribution in question summarizes the efforts of researchers and lecturers around the world who, in the areas of their expertise, hypothesize methods, tools and technologies for new knowledge and transparent communication. The research methodology is underpinned by the "Collaborative" type. The team avails itself of the following collaborations: Javier Julian Enriquez (Universitat Politècnica de València, Department of Applied Linguistics), Karen Alkoby Gallaudet (University of Washington); Mohammad Ayaz Ahmad (University of Tabuk, Saudi Arabia, Dip of Physics) Anastas Ivanov Ivanov Todor Kableshkov (School of Transport Higher, Dip of Mechanics), Roberto Capone (University of Salerno), Francesco Saverio Tortoriello (University of Salerno), Ritamaria Bucciarelli (University of Siena, DISFUCI), Marianna Greco (MIUR), Samuela Franceschini

(University of Venice), Giulia Savarese (University of Salerno), Raffaele Marcone (University of Salerno), Francesca Santoro (SIDELMED (Spa) and the contribution in question is the result of a joint work, for areas and international experiences The research focuses on mathematical models for the description of languages and creates new generation software for the construction of natural languages. The team entrusts the hypothesis and the entire validation process to the Popperian model in order to heuristically conjecture the moments of transformation, through which, over the centuries, the mathematical sciences, with the help of philosophy, psychology, computer science, economics, etc.. have raised the languages of Natural Language Generation (NLG) to formal models. In the first phase, it reinforces an epistemological framework, which entrusts the logical-mathematical reasoning of the LSP-ESP language to mental processes, following the models:- Scaramuzzino (2016) with an explanation of the phono-morpho-syntactic and semantic characteristics of the LSP languages; Bernardini (2010) which explains the question of replacing L1 in L2. The focal part of the research is probabilistic computation, according to Probabilistic latent semantic Analisis (Hofmann, 1999) - The work ends with the explanation of Digital intelligences. Our research question is: Is the ESP language a NLG language? We will try to give an answer keeping in mind that the sentence can be reconverted into spontaneous or pre-constituted algorithms and an algorithm can be considered a finite logical sequence of operations that is subject to mathematical laws. therefore, it can be reconverted into formal language and as such complies with the formal code LSP and ...

1. Analysis and understanding of LSP languages

It defines the specialized languages by Meeuwisse, A., Scaramuzzino, R. and Swärd, H. (2011). Sub codes, micro languages, scientific languages and technical languages are all special languages that can be classified strictly, if they have a specific lexicon that designates only concepts, which exist just in the relevant disciplinary field; and in the broadest sense, if they don't have a specific lexicon.

In addition to specific purposes in order to communicate clearly, the speaking communities that use these languages, are also easily identifiable socially because they operate in specific scientific or professional fields. The double advantage of clarity is therefore obtained by using the specific term and reducing the ambiguities deriving, for example, from the use of pronouns, it allows to be identified as belonging to the scientific or professional group that shares the particular language ... The specificity, therefore, of these languages is a high degree of specialization, belonging to a scientific or professional group and a well-defined situational context.

2. ESP: A language to be defined

ESP Definition: ESP is the teaching of English to students whose first language is not English but whom. ...English: esp. English: ESP Example sentences Trends ... ESP. is a written abbreviation for especially. The team defines ESP as an acronym mediated and reproducing a professional code NLG. An empirical research undertaken by Brindisi, G., Di Santo, L., Salvatore, S., Simonelli, M. A., Dante, V., Bernardini, M. G., ... & Di Carlo, L. (2018) on approaches for the study of the English language, at the Department of Interpreting and Translation of the University of Naples, shows that in Italian universities, due to the high frequency of foreign students in the future, according to Sheldrick, (2015), three forms of English could coexist: one that is spoken locally with dialectal influences, a national variety that will be spoken at work or at school, and finally the international standard English spoken by foreigners. Therefore, the professionals of writing must learn a foreign language with which to confront themselves, a major issue for their (future) work. The topic is open because the English language is currently the turbo engine of globalization, with the effect of obtaining specialized languages composed of sub-codes, i.e. a micro-language in the micro-language requires a new culture, for the teaching of the skills and training of the global citizen. The LSP language has become increasingly standardized, writing professionals must learn a foreign language to cope with a relevant topic, for their (future) work. In keeping with Bernardini, S., Ferraresi, A., Russo, M., Collard, C. and Defrancq, B. (2018), they are common, limiting the LSP label for both [general purpose English] and [English for specific purposes] and considers it inappropriate according to (Fig. 1)

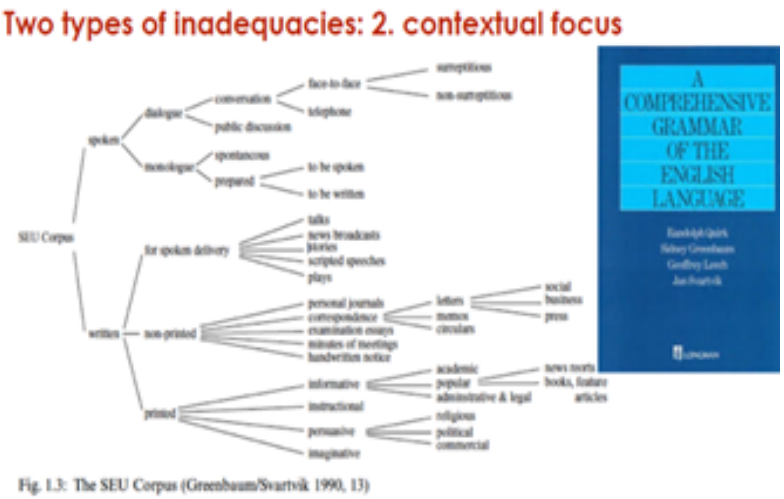

Fig. 1.3: The SEU Corpus (Greenbaum/Svartvik 1990, 13)

and syntagmatic regularities in the case that we consider that it is no longer possible to reduce a standard official language code to a list, a long list of individual variations such as Cohen, HY, Miller, C., Bitterman, KJ, Wall, NR, Hekking, B., Kessler, B., ... and Sinclair, DA (2004) point out:

> *It is no longer possible to imagine a sharp division between one type of patterning which behaves itself and conforms to broadly suitable rules, and another which is a long list of individual variations, and then to insist that they both create meaning at the same time. The patterns which are marginalized by our current attitudes include everything from collocation of all kinds, through Firth's colligations, to the conditioned probability of grammatical choices. This is a huge area of syntagmatic prospection. We identify structure like compounds, where the assumption is of a single choice, or idioms, although the precise identification of these is by no means clear-cut. The likelihood is of there being a continuum between occasional, quite independent choices and choices which are so heavily dependent on each other that they cannot be separated, and so constitute* in *practice a single choice*

Models that are marginalized by our present attitudes include everything from the placement of all types, through the colligations of Firth, to the conditional likelihood of grammatical choices. This is a huge area of syntagmatic exploration. We identify structures as compounds, where the assumption is of a single of choice, or idioms, even if the precise identification of these is not at all means "clear". It is likely that there is a continuum between occasional, fairly independent choices and choices that are so heavily dependent on each other that they cannot be separated, and thus would constitute a sort of single choice in practice as stated by Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y., ... & Hu, L. S. (2004). Prof. Bernardini does not clarify the continuum of demarcation, which is common to LSP-ESP, but which delegates to Biber (2012) the clarification on the common ground between corpus linguistics and ESP:

> Corpus-based studies generally have key research objectives:
>
> 1. Describe the variants and use of a linguistic word or structure, or
>
> 2) Describe the differences between texts and varieties of text, such as registers or dialects. The first objective concerns the classical studies of linguistic variation (for example, the choice between active and passive voice), while the second is used to investigate the linguistic variation of the text, contrasting the words and grammatical structures typical of different texts and varieties.

The team mentions a third no less important objective that can be defined by Silberztein, (2005).
3) The production of Language Resources for the automatic generation of natural language, i.e. the development of extensive language resources to develop NLG software in the real world. This research involves two communities on the one hand NLG researchers are generally more focused on specific technical issues for text generation, where good performance (e.g. recall and precision) is crucial, while linguists tend to focus on problems related to the development of exhaustive and precise resources that are primarily "neutral" towards any application of NLP (e.g., the development of a language resource for the real world) analysis or sentence generation), using various grammatical formalisms such as NooJ, Silberztein, M. D. (1994) TAG or HPSG.

3. Corpora for learning NLG languages

G. Aston argues that for language teaching one has to learn by means of corpora and then get the student to learn the phraseology of corpora discourse because:

> *There is substantial evidence that phraseology is primordial to fluency in speech production and reception, particularly in cognitively/affectively demanding contexts like interpreting. Yet most second language speakers have limited repertoires of phrase logical items, lacking knowledge of their lexico-grammatical, functional, and also of their prosodic aspects. Speech corpora which align transcripts with audio can readily be constructed from subtitled video materials, and learners can use these to view and hear concordance data. Examples are provided for phrase logical items documented in a one million-word corpus of talks from the TED – Ideas worth spreading (http://www.ted.com.),analysed using Word Smith Tools (Scott 2012). Activities for performance with and by learners are also suggested, aimed at increasing their phrase logical awareness and expanding their repertoires.*

It has been more than twenty years since Sinclair proposed his idea of "idiom principle" (1987:324; 1991. 101ff) as a central element of the production and understanding of language, based on the use of recovered and processed phrase logical elements. They can be composed not only of collocational regularity but can also be associated with specific semantic and pragmatic functions. Depending on how they are defined (as fixed multi-word sequences or as sequences that allow variability, and as having minimum or minimum frequencies of occurrence. In his dissertation, G. Aston reveals how it is necessary to acquire phraseology of corpora for the production of a professional communication. DDL techniques for teaching foreign language phraseology are in use for future interpreters. Guy Aston subscribes to the opinion that phraseology is the key to fluidity in both vocal production and reception, particularly in cognitively and affectively demanding contexts such as interpreting. The originality of this project lies in the use of a speech corpus for this purpose: a collection a million words of talks of the TED - Ideas worth Spreading site that aligns transcripts with audio/video files. With which model could we implement these approaches?

4. Filtering Methodologies

$$r_{u,i} = \bar{r_u} + k \sum_{u' \in U} \text{simil}(u, u')(r_{u',i} - \bar{r_{u'}})$$

where is the average rating of user *u* for all the items rated by *u*.

5. Discussion and results: Digital Intelligence Acro Word produced by Bucciarelli, P. Villari; F. Santoro; F. Terrone; M. Terrone; R. Marcone (© 2017)

As Javier Julian Enriquez points out: *"Consequently, based on both excellent theoretical and empirical research, this article illustrates the development and challenges of the applicative approach from a pedagogical and technological point of view, which can serve very useful purposes in the interdisciplinary field of computational linguistics."*. More to the point, this project addresses the application domain of computational codes. Having to define a set of possible words that represent the code and could be expressed in different languages, the application approach is based on Microsoft ® ACCESS. The result is the implementation of two main modules to assist the activities of generation and development of coding systems (DEFINER) and to process documents with code in order to obtain a fast and reliable translation (PARSER). The development of the project consists of a prototype stage of the two modules mentioned above, released in a single DB, necessary for the development of coding matrices and the basic characteristics of the parser. The prototype version contains all the basic performances identified in the project and related to the coding problem. The prototype will allow the development of technical and performance codes and tests of the PARSER, which refers to characteristics of the first final version, protection systems, interface, and user-friendliness aspects, etc. From the prototype version are evident the two areas of work, "Text" for the activities of drafting and translation of documents encoded on the basis of the coding system (formerly called PARSER module) and "Encoding" to allow the definition and maintenance of the coding system (activity of the module previously called DEFINER). For the writing and translation activity, the "vocabulary" option is available, which allows the writing of free text containing codes. Once the text has been prepared, the immediate translation of what has been written can be requested according to the coding system specified or available through "typewriters "In view of this, these empirical analyses, supported by tests of great value, may lead us to support and achieve these results: The results are the following software: Type-Race, for the automatic analysis of the text: it is a consultation of the linguistic corpus made in education almost to (program). Work tool, a database approach to manage lexicon construction and locate token modules in sign languages. Acro-word (human research): a multi-functional software used for lexical research and university teaching and consists of a.: a database (WT) transcribed into acronyms, called (linguistic corpus), for queries and analysis and "quantitative" reading of a linguistic corpus; text file, called (Human search converter) because it implies the presence of a digital operator who queries the database and uses the fixed parts in the test that goes to compose, use the acronyms in which they need an approval of the parts of the text and converts and transforms where it takes into account the specific parts of the text; text file called (beyond the language) for writing and simultaneous translation in real time on digital bar.

Acknowledgments

Thank you so much to the research team for their most important proposal, which mainly focuses on "the realization of a morphological, syntactic-semantic parser that is able to make structural choices based on customized codes" regarding the technologies of the editor texts. In this respect, the following article deals with a comprehensive introduction to state-of-the-art Text-To-Speech (TTS) synthesis by highlighting its Digital Signal Processing (DSP) and Natural Language Processing (NLP) components, which provide us with a helpful tool to learn a new language.

Riferimenti bibliografici

Aston, G. (2001). *Learning with corpora: An overview*. Learning with corpora, 7-45.

Bernardini, S. (2017), *Some reflection on empirical research approaches for the study English*. Department of Interpreting and translation.

Bernardini, S., Ferraresi, A., Russo, M., Collard, C., & Defrancq, B. (2018). *Building interpreting and intermodal corpora: a how-to for a formidable task. In Making way in corpus-based interpreting studies* (pp. 21-42). Springer, Singapore.

Biber, D. E. (2012). *Corpus-based and corpus-driven analyses of language variation and use. In The Oxford handbook of linguistic analysis*. Oxford University Press.

Brindisi, G., Di Santo, L., Salvatore, S., Simonelli, M. A., Dante, V., Bernardini, M. G., ... & Di Carlo, L. (2018). *La giustizia tra storia, genealogia e sociologia*.

Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y. ... & Hu, L. S. (2004).

Bucciarelli, R. (1998), *Lessico dell'italiano segnato: metodi,descrizioni, applicazioni*. Roma: Edizioni italiane.

Bucciarelli, R., Alkoby, K., Galdi, A., Greco, M., & Franceschini, S. (2018), *Heuristic filtering for the validation of scientific validation of emotional communication*. www. researchgate.com

Bucciarelli, R., Falco, V., & Savarese, G. (2019). *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.

Casiraghi, D., Sancassani, S., Tomasini, A., Corti, P., & Brambilla, F. (2019). *Il Modello "Pedagogy-Space-Technology" come Framework strategico a supporto dell'innovazione didattica*. In Reports on E-learning, Media and Education Meetings (Vol. 8, pp. 62-67).

Cefalato, P. (2004). *Critica della ragione postcoloniale. Verso una storia del presente in dissolvenza (Vol. 21)*. Meltemi Editore srl, Roma..

Chambers, A., Conacher, J., and Littlemore, J. (Eds.) (2004). *ICT and Language Learning: Integrating Pedagogy and Practice*. Birmingham: Birmingham University Press.

Chomsky, N. (1964), *Aspects of the Theory of Syntax. Massachusetts Inst of Tech Cambridge Research Lab Of Electro*nics.

Chomsky, N. (1986), *Tre modelli per la descrizione della lingua. Transazioni sull'IRE sulla teoria dell'informazione (2)*: 113-124. doi : 10.1109 / TIT.1956.1056813.

Chomsky, N., & Lightfoot, D. W. (2002), *Syntactic Structures*. Walter de Gruyter,

Churchman, C.W. (1971). *The Design of Inquiring System*. New York: Basic Books.

Cicalese, A. (1995), *L'analisi dei nomi operatori con il verbo fare. Tra sintassi e semantica. Descrizione e metodi di elaborazione automatica della lingua d'uso.* Loffredo, Napoli.

Cook, V.J. & Newson, M. (2007), *Chomsky's universal Grammar (an introduction),* Wiley-Blackwell.

Cohen, H. Y., Miller, C., Bitterman, K. J., Wall, N. R., Hekking, B., Kessler, B., ... & Sinclair, D. A. (2004). *Calorie restriction promotes mammalian cell survival by inducing the SIRT1 deacetylase*. science, 305(5682), 390-392.

D'Agostino, E., Elia, A., & Vietri, S. (2004), *Lexicon-grammar, electronic dictionaries and local grammars of italian. Lingvisticae Investigationes Supplementa*, 24, 125-136.

El-Behadli, A. F., Neger, E. N., Perrin, E. C., & Sheldrick, R. C. (2015). *Translations of developmental screening instruments: an evidence map of available research*. Journal of Developmental & Behavioral Pediatrics, 36(6), 471-483.

Elia, A., Landi, A., Bucciarelli, R. (2000), *Dalla grammatica al testo poetico*. Loffredo editore, Napoli.

Elia, A. (1984), *Le verbe italien: les complétives dans les phrases à un complément*. Schena; Nizet.

Elia, A., Martinelli, M., & d'Agostino, E. (1981), *Lessico e strutture sintattiche: introduzione alla sintassi del verbo italiano*. Liguori, Napoli.

Ferraresi, A., Bernardini, S., Picci, G., & Baroni, M. (2010). *Web corpora for bilingual lexicography: A pilot study of English/French collocation extraction and translation. Using Corpora in Contrastive*

and Translation Studies. Newcastle: Cambridge Scholars Publishing, 337-362.

Franceschini, S. (2019), Technologies and Didactics of Languages: Research-Experimentation-Productions.

Giunti, M. (1983), *Popper e Lakatos: due diverse giustificazioni del falsificazionismo*. Dimensioni, 28, 128-156.

Gross, M. (1986), Grammaire transformationnelle du français, Syntaxe du verbe. Cantilène, Paris.

Harris, Z. S. (1964), *Transformations in Linguistic Structure*. Proceedings of the American Philosophical Society 108:5, pp. 418-122.

Harris, Z. S. (1976), *Notes du cours de syntaxe*.

Ibrahim, A. H. (2003), *Le cadre du lexique-grammaire. Linx. Revue des linguistes de l'université Paris X Nanterre*, (48), 101-122).

Ibrahim, A. H., Dunn, P. F., & Brach, R. M. (2003), *Microparticles et achment from surfaces exposed to turbulent air flow: controller experiments and modeling. Journal of aerosol science*, 34(6), 765-782.

Ivory, J., & Gean, S. (1991). *A paradigmatic Analysis of Contemporary IT development.* European Journal of IT, 1(4), 249-272.

Jackendoff, R. S. (1972), *Semantic interpretation in generative grammar*.

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (1999), *El aprendizaje cooperativo en el aula*. https://www.researchgate.net/publication/265567256_El_aprendizaje_cooperativo_en_el_aula/comments?focusedCommentId=5bfc5905cfe4a764

Lamiroy, B. (2003), *In memoriam Maurice Gross*, Travaux de linguistique 46:1, pp. 145

Landi , E. Bucciarelli R. Landi , A. (2000), *Dalla grammatica al testo poetico: lezioni di linguistica*. Loffredo editore Napoli

Landi, A. (1999), *Saggio sulla varietà diamesica del Bellum Africum*. Hermes, 303-316.

Meeuwisse, A., 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.

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2002), *Computational and evolutionary aspects of language*. Nature, 417(6889), 611-617.

O'Malley, J.M. & Chamot, A. (1990), *Learning strategies in Second Language Acquisition*, Cambridge University Press.

Odlin, T. (1994), *Perspectives on Pedagogical Grammar*. (Cambridge Applied Linguistics).

Owen, C. (1993). *Corpus-Based Grammar and the Heineken Effect: Lexico-grammatical Description for Language Learners1*. Applied Linguistics, 14(2), 167-187.

Popper, K. R. (1963), *Conjectures and Refutations, Routledge and Kegan Paul*, London, trad. it. Congetture e confutazioni, Il Mulino, Bologna 1972.

Ricerca condotta in un progetto Collaborative dal titolo *Immersion in text typologies ed I.Ri.I.S,* Salerno https://www.researchgate.net/profile/Javier_Julian_Enriquez.

Rizzi, L. (Ed.). (2004), *The structure of cp and ip: The cartography of syntactic structures volume 2*: The cartography of syntactic structures (Vol. 2). Oxford University Press on Demand.

Ross, S. E., Ouyang, Z., Rajagopalan, S., & Bruns, T. M. (2018). *Evaluation of decoding algorithms for estimating bladder pressure from dorsal root ganglia neural recordings*. Annals of biomedical engineering, 46(2), 233-246.

Sabatini, F. (2007), *Il Sabatini Coletti*. Dizionario della lingua italiana 2008 con CD rom. RCS, Milano.

Sayad, A. (2002), *La doppia assenza*. Raffaello Cortina editore, Milano.

Scelzi, R. (2010), *Nel linguaggio specialistico*. Studi di glottodidattica, 3(3), 118-140.

Scaramuzzino, M. L. (2016). *Apprendimento cooperativo: sviluppare la collaborazione per migliorare il clima di classe* (Doctoral dissertation, Scuola universitaria professionale della Svizzera italiana (SUPSI)).