

How Does a Robot Speak? About the Man-Machine Verbal Interaction

Pierre-André BUVET¹, Bertrand FACHE² and Abdelhadi ROUAM³

¹ Paris 13 University, Théories Textes Numérique, 99 Avenue Jean Baptiste Clément, 93430 Villetaneuse, France
pierre-andre.buvet@univ-paris13.fr

² TEAMNET, 10 Rue Mercœur, 75011 Paris, France
b.fache@teamnet.fr

³ ONTOMANTICS, 959 Rue de la Bergeresse, 45160 Olivet, France
abdelhadi.rouam@ontomantics.com

Abstract. We present here an intelligent man-machine interface which will be integrated to social robots dedicated to the elderly. Firstly, we discuss the specificities of conversational agents, also called chatbots, by comparing them to the other man-machine interfaces. Then, we establish how linguistic intelligence contributes to the development of high quality conversational agents. We finish by explaining the functioning of a conversational agent based on a linguistic processing of non-structured information.

1 Introduction

We are witnessing the advent of robots in our society, as Issac Asimov the writer of science fiction predicted in the 50s [1]. If the existing robots are way behind to be considered as well performing as the ones in the science fiction literature, huge and remarkable progresses have been made in the field of robotics in the last few years, especially robots functioning as caregivers ‘social robots’. One of the characteristics of these robots is the integration of a system of verbal interactions between men and machines. Hence, this article highlights the characteristic of man-machine verbal interactions.

The senior market in rich countries is striking due to the increase rate of aging of the population which is linked to the growth of life expectancy. The need to accompany this phenomenon by practices, so called, aging well is a huge stake for these societies. As seniors advance in age, the question of losing their autonomy raises. The projections show an increasing gap between the non-autonomous people and the professionals who help them overcome their difficulties. Therefore, robots are presented as a sustainable solution to compensate the lack of staff [2].

Robots, which give people a helpful hand, are characterized by their aspect, mobility and their communicative capacity. The first characteristic is the aspect which is more or less humanoid. Besides NADINE, a social robot which resembles exactly to its designer Nadia Thalman, other robots have a general shape which is similar to the human body. It is just their sizes which look more like an infant than an adult, for example, the robots PEPPER, NAO and ROMEO of SoftBanks Robotics Company. Some robots have a more sophisticated shape which bears little resemblance to the human body; such as, the ELLI-Q Company robots. Sometimes, their upper part looks like a human head, for example, the robots CUTTI or the robot KURI of Mayfield Robotics Company. There exists, also, robots which have the appearance of an animal, either totally as the seal robot PARO of Inno3Med company, or partially as the bear robot ROBEAR of Riken company. The second characteristic concerns the mobility or the absence of mobility of the robot. The most humanoid robots have a manner of movement which simulates human walking, for example, the version ZORA of NAO. Less humanoid robots use wheels to move,

Copyright © 2019 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

for example the robots of Kompaï Company or the robot GIRAFF of GiraffPlus consortium. Other robots have mobile elements but they do not move, for example the robots ORIGIBOT of Origin Robotics Company. The third characteristic is the man-machine interaction. It can be focused on the non-verbal communication (tactile, auditory or kinesis), in order to create emotions for users with cognitive difficulties, which is the case with the therapeutic robot PARO. The use of a screen integrated in the robot is another possibility, for example the range of robots of Padbot Company. The interaction with a screen may be tactile or vocal. At last, robots allow the verbal interactions through answering questions and obeying orders. Thus, the version ZORA of NAO interacts with its users through conversations.

Robots, which give a helpful hand to old individuals, have many functions which mobilize specific technologies. The function of care is the oldest one. It consists of measuring the physiological state of a person by pressing on objects connected to a platform dedicated to medical supervision. When the signals detect abnormalities, an alert is launched. Another function is the prevention of individuals' falls. It is applicable to robots which can be called smart walkers. They help people with balance troubles to move without risks. Developed computer systems manage the movements of these robots on the basis of the users' indications and information about the spatial environment obtained from sensors. Another function of robots is the detection of falls. They, also focus on the geo-location of sensors. They inform a system which measures the position of the user and launches an alert when this position is out of the security norms. Tele-presence is a broader function. On the one hand, it allows watching after a senior from a distance by moving the robot in the senior's environment and launching an alert in case of a problem. On the other hand, it communicates with the users. It involves a screen, which uses the videoconference and a system of remote control from a distance to move the screen mount. The function of sociability can be the fruit of tele-presence since it includes interactions between the user and another person. More broadly, the function of sociability concerns the companion robots, which are developed to compensate the emotional and social isolation of elderly. The function of helpfulness is consubstantial with robots since they perform tasks they were designed for. The last two functions, among others, focus on the technologies of chatbot type when it comes to verbal interactions. The quality of the chatbot is a crucial factor for the functions to work effectively.

For now, we are discussing only about chatbots. After placing the conversational agents¹ at the man-machine interface, we precise the contribution of linguistic intelligence to the development of an elaborated chatbot, then we explain how a conversational agent, based on a linguistic processing of a non-structured information, functions.

2 About the conversational agents

Digital technologies process the human language with its written and oral forms. These digital technologies either interpret it or automatically reproduce it. They are dedicated either exclusively to the first part of processing, for example the analysis of emotions, or specifically to the second part, for example automated journalism, or to both parts simultaneously. Digital technologies dedicated to the simultaneous processing of both parts are regularly man-machine interfaces known as 'chatbot' or 'conversational agents'. Chatbot, which is a compound English word, consists of the word chat and robot. A chatbot is presented in the form of an API which takes, as input, a sound or a text file and produces, as output, a sound or text file. This is followed by means of communication with the machine either in writing or speaking. The chatbot functions through the understanding of a message sent by a human being, a request, and through the formulation of another message, linked to the initial one, by the machine, a response.

Forms, interactive menus, command prompts, etc. allow the user to give instructions to a computer. Such modes of communication are, sometimes, insufficient to meet the user's exact needs due to their rigid relativity.

¹The terms conversational agent and chatbot are synonyms.

The man-machine dialogue is conceived as a means to deal with these inadequacies. It consists of a system that allows an interaction, which is less or more narrow, between a human and a system. Its implementation is difficult because of the complexity of the language.

Therefore, the man-machine dialogue is a man-machine interface type as the language is a means to exchange the data. The word *dialogue*² indicates “different shapes of interviews between two persons [3]”. The denotation *dialogue* has to be distinguished from *conversation* and *discussion*, because a dialogue has specific characteristics: it is structured to contain a beginning, a flow and an end; it involves codified turn taking; it needs a temporal isotopy, it means that there is no time lag in the verbal interactions [4]. Spatial isotopy is not a dialogue condition: we can make a dialogue with a person at a distance, for example through a phone call. Speaking and writing are the two possible ways of a dialogue. As a result, instant messaging permits people to have a dialogue. However, a correspondence and an exchange of mails don't belong to the category ‘dialogue’ because there is a time lag in the exchanges. The dialogue is attributed to two speakers, Speaker1 and Speaker2, who obey the principle of turn taking as the role of the speaker and the interlocutor is successively attributed to Speaker1 and Speaker2 [5]. In the case of the man-machine dialogue, one of the two speakers is a machine and the natural language is substituted with data processing codes. This substitution involves bilateral exchanges into the natural language. However, it is a question of either a voice user interface or a user-friendly interface with speech. Systems of the former type understand the human language by responding to oral injunctions. Systems of the latter type react through pronouncing instructions. The man-machine dialogue has been under research in natural language processing and artificial intelligence for more than thirty years [6]. It is a central sector of activity in Language Industry. In fact, man-machine dialogue systems are distinguished by their mode of running which can be more or less by the machine [7].

When the restrictions on use are striking, the verbal interactions have a big knowledge gap, as the GUS system that is dedicated to online air ticket sales [8]. In such systems, the notion of script has a leading role because it determines the nature of interactions and their organization mode. Outgoing messages are those of the artificial speaker; they are preset according to the understanding of the user's intention. The incoming messages are those of the human speaker, yet it is very often proven that “every man-machine dialogue system has its limits, and, quickly, the user notices that the machine has understanding limits [3]”.

Communication systems in natural language of the first generation are fundamentally dedicated to database queries or to very specific tasks. They permit to formulate instructions in natural language and answer them in natural language. Interaction systems in natural language of the second generation are more developed. They can deal with linguistic phenomena that impede the system effectiveness. To detect the human speaker's intention, it is necessary, in particular, to deal with the inference, the reference and the meta-dialogue. The processing of the implicit meaning consists of highlighting the implicit consequences of a statement [9]. For example, the fact of saying *the door is open* does not produce the same reactions according to contexts. This can mean ‘the door must be closed’, for example because of a draught or on the contrary, ‘the door must be open’, because somebody is coming. The processing of the inference consists of implementing extra-linguistic knowledge and tools capable of detecting the implicit meaning of a speech in a man-machine dialogue system [10]. The question of reference is crucial in natural dialogue processing. First, there is the answer to the question of first and second personal pronouns that change the referent in turn taking and depend on the situation of utterance [11]. Then, there is the analysis of the references set which concerns the third personal pronoun, nominal groups and proper nouns [12]. They function differently; it depends on either their reference functions in an anaphoric or a deictic manner. In the first case, their reference is found in the left context (it is found in the right context if the function is cataphoric). In the second case, it is to be found in the utterance situation. The processing of the reference needs to implement the same resources as before if it doesn't require to use tools capable of identifying the referents as well [13]. The meta-dialogue concept refers to the phatic function of language. Its processing “consists of specific sub-dialogues

²There is no special use of the word ‘dialogue’, not the one that is used in the theatre repertoire.

of repeat requests (followed or not by particular instructions), of confirmation or validation requests in case of ambiguity, of processing of the interlocutor questions on the results of understanding or recognition, of put on hold and of the dialogue maintain and revival [7]”. Man-machine dialogue systems of the third generation use works based on machine learning [14]. They permit, in very constrained operational framework, to perform more effectively certain tasks carried out in other system types [15].

There are many types of chatbots. We can subdivide them into two categories [16]: a) chatbots which make illusion but are proven ineffective in use; b) chatbots more performing, namely intelligent chatbots. Category a) chatbots can simply make two or three general exchanges, and then show their inability to process the information. Category b) chatbots are able to process the language information when the latter is applied to a limited sector. There are many implemented solutions to develop chatbots: 1) non-linguistic solutions; 2) weakly linguistic solutions; 3) quasi-linguistic solutions; 4) linguistic solutions; 5) strongly linguistic solutions.

Whatever the solutions, they can integrate technologies namely, on the one hand, speech-to-text, and, on the other hand, speech synthesis. In the first case, it is about converting automatically an oral message into a written one. In the second case, it is the contrary; a written message is automatically converted into an oral one. These two technological aspects are not presented here. The solution types 1 to 4, generally, correspond to a question-and-answer system, namely a question is asked by a human being and the answer is given by a machine. Type 5 solutions are considered to be more than a question-and-answer system, since it is about producing deeper verbal interactions between two human beings as one is a human being and the other is a machine.

Type 1 solutions are trivial because they are based on the preparation of ready sentences that correspond to requests which, once recognized, are associated with beforehand recorded answers. The first chatbots designed on this principle date from about fifty years ago [17]. Type 2 solutions are more developed since they use keywords that automatically identify in requests in order to associate them to concepts, which in their turn are associated with already made answers. A big number of chatbots are based on this principle. They can be reinforced by statistical tools that use the identified request-answer association [18]. Type 3 solutions are namely quasi-linguistic because they include linguistic analysis for the chatbot component which processes the understanding of the request, *cf. infra*. However, it is not the case for the formulation of the answer; it is an already made sentence. Type 4 solutions integrate the request analysis part of type 3 solutions and rely on the natural language generation, *cf. infra*. This technology permits to give a similar response in the content, but presents various shapes. Type 5 solutions have to allow the robot to simulate a conversation to the maximum by relying on verbal interaction scripts. It is to be conceived.

3 Linguistic Intelligence Contribution³

By definition, an intelligent chatbot automatically understands the incoming messages (a question) and it associates them automatically to outgoing messages (a response). It aims at simulating verbal interactions between two humans by replacing one of the speakers (typically, the one who gives answers). The chatbot performance depends on its ability to interpret the incoming messages and produce relevant informative outgoing messages. In addition, its performance is, also, conditioned by the speed of exchanging information between the user and the machine.

In order to improve the performance of a chatbot, it has to include the linguistic processing of incoming and outgoing messages. The processing of the incoming messages differs from the one of the outgoing messages since in the first case, it relies on natural language understanding of texts and in the second case, it relies on the natural language generation; that is to say two different technologies of natural language processing [7]. So far, natural language understanding and natural language generation technologies can share the same data model to process

³About the concept of linguistic intelligence and its relationship with the concept of artificial intelligence, *cf.* P.-A. Buvet. sous presse, “*Linguistique et intelligence*” in *Linguistique et ...* Peter Lang

either the incoming or the outgoing messages. And even more, this has to involve a device associating an incoming message to an outgoing one.

In order to facilitate the flow exchange in the information processing set, there has to be a formal and common representation of the language data. Here, the propositional contents presentation of incoming and outgoing messages is used, in terms of predicate-argument structure to understand and generate messages in a homogeneous manner. The predicate-argument structure concept relies on Zellig S. Harris works [19], Maurice Gross [20], Pierre-André Buvet [21], Salah Mejri [22] and Robert Martin [23]. These works are inspired by the predicate logic suggested by Gottlob Frege [24].

In this theory, the concept of predicate is defined as a relationship in which the propositions, in its logical term, are analyzed as a link between entities, named arguments. This modeling differs from the Aristotelian modeling and gives place to the functional representation of the propositional content⁴: proposition => prédicat (argument)_i⁵
6

Applied to language facts, the calculation of predicates leads to associate the contents of an utterance to predicate-argument structure. For example, the utterance *Le médecin soigne un patient* (*The doctor treats a patient*) is metalinguistically represented by the following predicate-argument structure: soigner (médecin,patient)⁷. Utterances are distinguished by the fact of having simple or complex propositional content. In the second case, nested predicates characterize the utterances. For example, in the utterance *Le médecin affirme au patient qu'il est guéri*, (*The doctor tells the patient that he is cured*) the propositional content integrates another propositional content. Hence, the utterance is represented metalinguistically by a predicate-argument structure mentioning a predicate nested in another predicate: affirmer(médecin,patient, guéri(patient))⁸.

A more sophisticated abstraction layer is also possible by replacing lemmatized forms of lexical units by their semantic class. Thus, both of the utterances above have the following metalinguistic representation:

```
SOIN(HUMAIN1,HUMAIN2)  
AFFIRMATION(HUMAIN1,HUMAIN2,GUERISON  
(HUMAIN2))9
```

To process the information of the incoming and outgoing messages, the choice of linguistic formalism is crucial. Thus, it is necessary to precise the type of the theoretical approach which underlies it. Theories about language are distinguished by their privileged subject of study. Therefore, to study the language facts, morphological theories have as an entry point the morphology, for example the theory named construction morphology [25]. It is the same thing for the syntactic theories, semantic theories, pragmatic theories and the theory of enunciation compared, respectively, to syntax, semantics, enunciation and pragmatics. Lexical theories assign a central place to lexicon in the linguistic analyses they produce. It is the case of those which rely on predicate-argument structures to represent language facts.

The effectiveness of a chatbot is based on the understanding of the incoming message and its adequacy with the outgoing message. From this point of view, the predicate-argument structures contribute to the association of an incoming message and an outgoing message in the chain of processing information by allowing the connection between the outcome of the interpretation of the incoming message and the beginning of the production of the outgoing message. The solution presented here is based on natural language understanding technology and a natural language generation technology, *cf. infra*. The junction point between these two technologies is the arrival

⁴ The representation is namely functional because the predicate corresponds to an algebraic function, the argument is a variable.

⁵ The letter *i* indicates the number of arguments in relation to a predicate, namely arity.

⁶ Proposition => predicate (argument)_i

⁷ To treat (doctor, patient)

⁸ To tell (doctor, patient, cure (patient)).

⁹ TREATMENT(HUMAN1,HUMAN2).

TELL(HUMAN1,HUMAN2, HEALING(HUMAN2)).

point of the first technology in the form of predicate-argument structure and the departure point of the second technology, also, in the form of predicate-argument structure. When the incoming message expresses a need, its metalinguistic representation is incomplete, namely one of its components is missing. The need expressed initially is satisfied in the outgoing message when the metalinguistic representation, which produces it, is saturated, namely the missing component of the other presentation is specified. The other elements of the two metalinguistic representations being similar are linked in a way that an association exists between the incoming message and the outgoing message. From this point of view, the predicate-argument structures are adapted, as metalinguistic representations, in order to make an association between the incoming message and the outgoing message.

The data model used to develop a chatbot grants an important role to the concepts of predicate and argument and to the relationship between them to automatically analyze language facts. The latter uses linguistic resources which correspond to local grammars, dictionaries and basic rules, *cf. infra*. The homogenization of the integration and the use of these resources in the system, which makes the conversational agent functions, creates the first constraint for the flow of information processing. Especially as, the processes of understanding and generating messages have opposite objectives. In the first case, the input is a text and the output is a metalinguistic representation. It is the opposite in the second case. The input is a metalinguistic representation and the output is a text. Moreover, it is necessary to associate the output of the natural language understanding to the input of the natural language generation. A human being conversation simulation represents a second constraint. From this point of view, the quality and relevance of the outgoing message is essential as well as the rapidity of the processing; the time of reply has to be about a second. Finally, the third constraint concerns the robustness of the system and its capacity to adapt to all kinds of computer science environment.

The chatbot relevance rate has been the subject of a quantitative and qualitative evaluation through the evaluation of two system components: on the one hand, the one which processes the understanding of the incoming messages, on the other hand, the one which is in charge of the outgoing messages. In both evaluations, we have to measure the noise, non-relevant information presented by the system as relevant, and silence, relevant information not selected by the system [26]. The first evaluation relies on the one used by labeling systems; it concerns the comparison of beforehand information identified by at least a person with the identified and qualified information by the system. It consists of manually labeling an incoming message corpus [7] and compares the manually labeling with the labeling of the same corpus by the semantic analysis engine. In that respect, noise and silence are measured in terms of precision and recall rates¹⁰. The results are as the following: Precision rate: 92.1 / Recall rate: 94.8.

The second evaluation consists of checking if the results given by the solutions are in accordance or not with the expressed needs¹¹. The evaluation protocol requires a person who doesn't know the given results by the chatbot in order for the comparison to be independent. In addition, the evaluation consists of manually develop an outgoing messages corpus and compare it with incoming messages produced by the chatbot. It appears that 11% of the manually obtained data are missing, but only 4% of the automatically given data are not in the reference corpus. The failures of the automatic generation module are due to questions of synonymy which appears inadequately processed until now.

¹⁰ It is useful to suggest the following formulations to calculate them: $Precision = \frac{true\ positive}{true\ positive + false\ positive}$ / $Recall = \frac{true\ positive}{true\ positive + false\ negative}$. The terms 'true positive' and 'false negative' indicate respectively the non-relevant information but identified as relevant and the relevant information but not identified as it is.

¹¹ The second evaluation differs from the first one because the generation process does not match with the beforehand manually made tag process. The control over the data, that the natural generation method involves, confirms that the noise concept has no point in the evaluation.

4 Operating Mode of Conversational Agents

We present the architecture of the system which makes the chatbot functions.

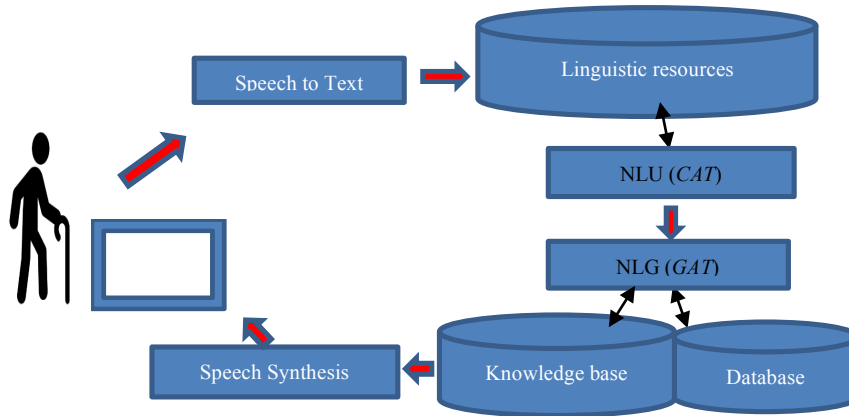


Fig. 1. The architecture of the system

API (Application Programming Interface) is the interface which permits the interactions between men and machines. Speech-to-text and speech synthesis modules have a role of transforming sound files into text files and another text file into a sound file. The Natural Language Understanding (NLU) module includes a semantic analysis engine which uses linguistic resources in the form of electronic dictionaries and local grammars [27]. The semantic engine's task is to replace the incoming message, in the form of text file, into a symbolic representation. This representation is given to the Natural Language Generation (NLG) module which associates the former representation to another symbolic representation from which an outgoing message will be produced in the form of text file.

The symbolic representation used in the system are metalinguistic representations, more precisely, it concerns semantic representations by making linguistic functioning mode of utterances. Incoming messages processed by the conversational agent have information requests or action requests. Information requests are introduced either in the form of questions (for example, *Que mangeons nous ce soir?* (*What do we eat this evening?*)) or in the form of requests (for example, *Je voudrais savoir ce que nous mangeons ce soir* (*I would like to know what we eat this evening.*)). Action requests are formulated in the form of orders (*Monte le chauffage* (*turn up the heating*)) or observations (*J'ai froid* (*I am cold*)). In the case of orders, the request is often formulated explicitly. In the case of observations, the request is implicit in a way that there has to be an inference in the observation about the need to be satisfied. Replies to action requests have two dimensions: extra-linguistic dimension, it concerns a concrete reply to the request in an appropriate way; and a linguistic dimension, it has to be known that the action requested is executed or about to be executed. Here, the extra-linguistic dimension is not developed.

When dealing with information or action requests, their metalinguistic representations include predicate-argument structures, cf. *supra*, in order to homogenize the process of information. Incoming messages are different in terms of the request type. In information requests, the human speaker's intention expresses an informative need which has to be satisfied. Also, the metalinguistic representation of an incoming message, henceforth RM_i (RM_e) has to specify this need in a way that can be associated with a metalinguistic representation of an outgoing message, henceforth RM_o (RM_s), relying on the fact that this representation can satisfy its need. Metalinguistic representations expressed, in terms of predicate-argument structures as functional representations, correspond to the following formula according to if the predicate is: a monadic predicate (*prédicat monadique*), namely it takes only

one argument, a unary operator (RM1 = PREDICAT (ARGUMENT1))¹²; a dyadic predicate (*prédicats dyadique*), namely it takes two arguments, binary operator (RM2 = PREDICAT (ARGUMENT1, ARGUMENT2))¹³; a triadic predicate (*prédicats triadique*), namely it takes three arguments, ternary operator (RM3 = PREDICAT (ARGUMENT1, ARGUMENT2, ARGUMENT3))¹⁴

The hypothesis has been made on the basis of the R_{Mi} which consists of merely one unknown element, a variable or a function, which is associated with the R_{Mo} from their other common elements which suggest the unknown element. Hence, the information request *Qu'est-ce qu'il a comme repas? (What do you have for lunch?)* has the following R_{Mi} MENU (X) and it is associated with the R_{Mo} MENU(%liste_plat%)¹⁵ which is associated with the R_{Mo} MENU (%liste_plat%) because the predicate MENU appears in both representations in a way that the argument %liste_plat% is the object of the question represented by X. It is the same case with the information request *Est-ce que mon fils m'a appelé? (Did my son call me?)*, it has the following R_{Mi} OUI_NON=X (TELEPHONE (FILS (INTERLOCUTEURh), INTERLOCUTEURh))¹⁶ and it is associated, according to the context, to either the R_{Mo} OUI (TELEPHONE(FILS(INTERLOCUTEURh), INTERLOCUTEURh))¹⁷ or to the R_{Mo} NON(TELEPHONE (FILS (INTERLOCUTEURh), INTERLOCUTEURh))¹⁸ because (TELEPHONE (FILS (INTERLOCUTEURh), INTERLOCUTEURh))¹⁹ appears in both of the representations in a way that the predicate OUI (yes) or NON (no) is the object of the question represented by the predicate OUI_NON=X.

In action requests, the human speaker's intention corresponds to a concrete need that has to be satisfied. An incoming message in a form of an order is sufficiently explicit to be returned to a R_{Mi}. However, when it is in a form of an observation, its intention is indirectly accessible to the interlocutor; there must be an inference of the semantic contents to express other semantic contents. The latter has to be added to the R_{Mi} in a form of a predicate-argument structure to the predicate-argument structure which has a link with the first semantic contents. There are inferences' rules used to result in such R_{Mi}. The association of these representations to the R_{Mo} is different from the previous one. It concerns the identification, among the second ones, of similar predicate-argument structures with the first ones and concatenates them with a predicate structure to execute the initially formulated request. For example, the action request *Il fait sombre? (Is it dark?)* has the following R_{Mi} FAIBLE LUMINOSITE (DEIXIS) & ORDRE (INTERLOCUTEURh, INTERLOCUTEURr, ALLUMAGE (LUMIERE)) and it is associated with the R_{Mo} FAIBLE LUMINOSITE(DEIXIS) & ORDRE (INTERLOCUTEURh, INTERLOCUTEURr, ALLUMAGE (INTERLOCUTEURr, LUMIERE)) & INFORMATION (INTERLOCUTEURr, INTERLOCUTEURh, EXECUTION ORDRE (INTERLOCUTEURh, INTERLOCUTEURr, ALLUMAGE (INTERLOCUTEURr, LUMIERE))).²⁰ The results R_{Mo}, for example to the following utterance is: *Comme il fait sombre, vous souhaitez que j'allume la lumière. C'est fait (Because it is dark, you want me to turn on the lights. It is done).*

The expressive language of semantic contents, an intention, comes from a huge diversity at the morphological, syntactical and lexical levels [28]. This transformation happens at the level of the text understanding module

¹²(MR1 = PREDICATE (ARGUMENT1))

¹³(MR2 = PREDICATE (ARGUMENT1, ARGUMENT2))

¹⁴(MR3 = PREDICATE (ARGUMENT1, ARGUMENT2, ARGUMENT3))

¹⁵(%list-dish%)

¹⁶YES-NO=X (TELEPHONE (SON (SPEAKERh), SPEAKERh))

¹⁷YES (TELEPHONE (SON (SPEAKERh), SPEAKERh))

¹⁸NO (TELEPHONE (SON (SPEAKERh), SPEAKERh))

¹⁹(TELEPHONE (SON (SPEAKERh), SPEAKERh))

²⁰LOW LIGHT (DEIXIS) & ORDER (INTERLOCUTORm, INTERLOCUTORr, TURN ON (LIGHT)) and it is associated with the R_{Mo} LOW LIGHT (DEIXIS) & ORDER (INTERLOCUTORh, INTERLOCUTORr, TURN ON (INTERLOCUTORr, LIGHT)) & INFORMATION (INTERLOCUTORr, INTERLOCUTORh, ORDER EXECUTION (INTERLOCUTORh, INTERLOCUTORr, TURN ON (INTERLOCUTORr, LIGHT))).

through two steps: 1) data extraction through the identification of their linguistic form; 2) qualification and interpretation of the data by associating them to semantic meta-information. To interpret these incoming messages, there has to be a semantic analysis engine which simulates three linguistic capacities made by humans. The first capacity is the lexical capacity, namely the memorization of simple or compound words. The second capacity is the structural capacity which concerns the morphological, syntactical and semantic levels of the language. The third capacity is the combinatorial capacity, namely to express the same propositional content in all kinds of ways. The simulation of these three capacities needs Information Technology tools and linguistic resources to identify the information and qualify it with semantic tags, *cf. infra*. The choice of adding meta-information to texts in terms of predicate-argument structure relies on the identified lexical units' properties. They have to be sufficiently described in the linguistic resources for the semantic labeling to operate.

The transformation of a RMo into an outgoing message is another difficulty. The choice has been made from a process which relies on conceptual dictionaries, namely dictionaries whose macrostructure is formed of concepts and the microstructure of lexical units associated with each concept. There are two conceptual dictionaries: the one of predicates and the one of arguments. The RMo are in form of predicate-argument structures as the predicate corresponds to the function and the arguments correspond to its variables. The predicates and their arguments are symbolized by their semantic class. Though the pragmatic adjustments, for example *Lundi (Monday)* when the incoming message is *Quel jour sommes-nous? (What day are we?)*, the semantic class is related to the lexical units which they characterize. For example, the semantic class METEO_NEIGE²¹ characterizes the verb *neiger (to snow)*, the noun *neige (snow)* and the adjective *neigeux (snowy)*. The lexicon instances of a predicate, unlike the ones of arguments, have the particularity of being characterized by standard constructions. For example, the construction X0:PRONOM V²² characterizes the verbal predicate *neiger (to snow) (il neige) (it snows)*, the constructions X0:PRONOM y avoir DU N and X0:GROUPE_NOMINAL être à LE N²³ characterize the nominal predicate *neige (il y a de la neige et le temps est à la neige) (snow) (there is snow and weather forecast calls for snow))* and the construction X0:PRONOM+GROUPE_NOMINAL être A²⁴ characterize the adjective predicate *neigeux (snowy)*. The unique argument, namely X0, is semantically specified at the distributional semantic level; it is the subject in all the constructions as deictic, nominal group or both of them. There are reconstructions associated with constructions, namely non-standard constructions or the component order is modified in a way that the syntactic modification does not involve the major semantic modification. For example, the construction X0:GROUPE_NOMINAL être à LE N²⁵ (*Le temps est à la neige) (weather forecast calls for snow)* has as a possible reconstruction, ce être à LE N que être X0:GROUPE_NOMINAL²⁶ (*C'est à la neige qu'est le temps) (it is called for snow by weather forecast)*. The juxtaposition of these different meta-information gives rise to sentence patterns in the instantiation phase of the lexical units in the constructions, namely when the major components are replaced by a beforehand specified vocabulary. Morphosyntactic resources are, then, used to apply formal grammar to sentence patterns and produce well-formed utterances. Linguistic descriptions are formalized in electronic dictionaries and rule bases. Their good quality and systemicity are fundamental for the well-functioning of the Natural Language Generation module. The latter simulates the human language by suggesting all kinds of utterances relevant to the same semantic contents. The objective is met because the lexical and syntactic variety is taken into consideration by the data model implemented by the system.

²¹WHEATHER FORECAST_SNOW

²²X0: PRONOUN V

²³X0: PRONOUN have of N and X0: NOMINAL_GROUP to be at THE N

²⁴X0: ARTICLE+NOMINAL_GROUP to be A

²⁵X0: NOMINAL_GROUP to be at THE N

²⁶This to be at THE N that to be X0: NOMINAL_GROUP

In the Natural Language Understanding module, the incoming messages are enhanced with semantic tags, dictionaries and *local grammars*. Below, we introduce an excerpt of one of these dictionaries, the body-parts dictionary. This dictionary contributes to the identification and qualification of information relevant to the health of the conversational agent users.

```
annulaire,.N+H_PARTIE_CORPS
(ring finger,. N+H_BODY_PART)
annulaires,annulaire.N+H_PARTIE_CORPS
(ring fingers,ring finger. N+H_BODY_PART)
articulation,.N+H_PARTIE_CORPS
(joint,. N+H_BODY_PART)
articulations,articulation.N+H_PARTIE_CORPS
(joints,joint. N+H_BODY_PART)
avant bras,avant-bras.N+H_PARTIE_CORPS
(forearms,forearm. N+H_BODY_PART)
avant-bras,.N+H_PARTIE_CORPS
(forearm,. N+H_BODY_PART)
auriculaire,.N+H_PARTIE_CORPS
(auricular,. N+H_BODY_PART)
auriculaires,auriculaire .N+H_PARTIE_CORPS
(auricular,auricular. N+H_BODY_PART)
bouches,.N+H_PARTIE_CORPS (mouths,. N+H_BODY_PART)
```

Fig. 2. A dictionary relevant to the health of the conversational agent users

The information on the left side of the full stop is of a linguistic nature and the one on the right side of the full stop is of a meta-linguistic nature. The macrostructure of the dictionary is made of lexical units constituting the dictionary. The microstructure is constituted of a lexical unit followed by a comma and its lemmatized form if the entry is an alternative. Next, there is a full stop which is mandatory followed by a grammatical category (N is the code for noun) and by meta-information of a semantic nature, in this case, the code H_PARTIE_CORPS (H_BODY_PART) is relative to body part hyperclass. This meta-information may be called from a local grammar in a way that all the dictionary items are recognized.

Below, we introduce one of these local grammars, the local grammar namely PROBLEME_SANTE (HEALTH_PROBLEM).

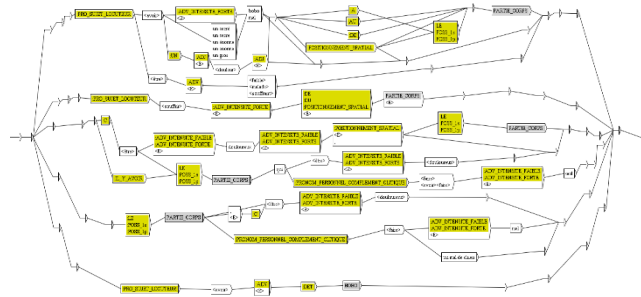


Fig. 3. Local grammar (HEALTH_PROBLEM)

Local grammars are formal representations of contextual elements. Regular grammars used by a parser which constitute the simplest grammar class in Chomsky hierarchy [29]. The representation of a finite state automaton is the one of an oriented graph (namely with a departure point and an arrival point) in which the windows are the states and the arrows are the transitions between these states. They correspond to formalized descriptions of syntax of a semantic class or a grammatical class. They are implemented in forms of automata which run through texts in order to identify and qualify the information. Local grammars use electronic dictionaries which are formal representations of lexicographical data. They are introduced in forms of graphs involving a start node and an end node and information nodes of lexical or morphological nature, the nodes linked make different possible combinations, regressions and repetitive structures. A graph can call other graphs which makes its combinatorics even more powerful. In computer science, graphs correspond to either finite state automata or finite state transducers. Finite state automata allow the identification of information by running linearly through texts and by reporting each connection between the text and one of its graphic paths. Finite state transducers integrate the first ones; they identify information according to the same principles and qualify them by inserting new data.

PROBLEME_SANTE (HEALTH_PROBLEM) graph calls other graphs and one of them is the PARTIE_CORPS (BODY_PART) graph which represents the dictionary PARTIE_CORPS. It allows the identification of all kinds of utterances which have a relationship with a health problem related to the human speaker, for example, *J'ai mal à la poitrine* or *Mon ventre me fait mal* (*I have pain in my chest* or *My stomach hurts*). It qualifies them, by adding a tag, in a predicate-argument structure form PROBLEME_SANTE (LOCUTEUR_HUMAIN)²⁷.

The information specified in the utterance is not always explicit. For example, in the use of a chatbot context, reporting a physical pain implies the implicit information '*j'ai besoin d'être soigné*' ('I need to be treated'). It follows that the tag PROBLEME_SANTE (LOCUTEUR_HUMAIN) is interpreted as an action request and it has to be completed with the tag APPEL (LOCUTEUR_MACHINE, PERSONNEL_SOIGNANT)²⁸. Such transformations happen when the RMo are given to the Natural Language Generation module and they are processed by inference rules.

In the Natural Language Generation module, the semantic representation of the outgoing message is a predicate-argument structure, *cf. supra*. From this representation, it is possible to generate a well formulated utterance by applying well ranked rules which consider: semantic properties of a distributional and lexical nature; syntactic properties related to the predicate constructions and the positioning of the lexical material, specified in a previous rule, in these constructions (a procedure named *instanciation* (instantiation)); enunciative properties relative to the actualization of predicates and arguments (a procedure named *actualisation* (actualization)); morphosyntactic

²⁷ HEALTH_PROBLEM (HUMAN_SPEAKER).

²⁸ CALL (MACHINE_SPEAKER, NURSING_STAFF).

properties which give rise to verb conjugation and coordinating conjunction rules with adjectives and nouns; pragmatic properties which narrow the field of possibilities and take the extra-linguistic context into consideration.

The semantic, syntactic and enunciative properties are listed in a conceptual dictionary. This dictionary is transferred to the database. The fourth category of properties is about the morphosyntactic dictionary, created as part of the project. This second dictionary is also inserted in the database. The fifth category of properties is covered by the knowledge base.

These linguistic and extra-linguistic resources are used by the natural generation module through the implementation of a set of rules which use linguistic resources inserted in the database. Below, we list the different rules used. The order of their presentation corresponds to their order in the passages before.

Predicate identification rule

Semantic properties of a distributional nature define propositional contents as the predicates and their arguments which are specified in a functional representation form as a semantic value. For example, utterances such as *Il y a du brouillard, C'est brouillardoux* and *Le temps est au brouillard* (*There is fog, It is foggy* and *The weather forecast calls for fog*) share the same semantic contents represented as the following: METEO_BROUILLARD (DEIXIS+METEO)²⁹. The first rule used by the natural generation module relies on this type of meta-linguistic description.

The departure point of the generation module is a predicate-argument structure. Its functional representation indicates if we are dealing with a simple predication or a complex predication. The second predication is distinguished from the first one by the fact that it incorporates at least another simple or complex predication, in its argument domain. The functional representation of a complex predication involves at least a double parenthesis, which allows its automatic identification. For example, the utterance *Le médecin est absent* (*The doctor is absent*) is analyzed as a simple predication which corresponds to the following predicate-argument structure: 1) (ABSENCE (HUMAIN:MEDECIN))³⁰.

Furthermore, an utterance as: *Le médecin a déclaré aux patients qu'il était absent* (*The doctor announced to his patients that he was absent*) is analyzed as a complex predication which corresponds to the following predicate-argument structure: 2) (DECLARATION(HUMAIN:MEDECIN,HUMAIN:PATIENT,ABSENCE(HUMAIN:MEDECIN)))³¹.

The end of the representation 1) involves one closing parenthesis while the representation 2) involves two closing parentheses. The number of the closing parentheses indicates the type of the predication. If it is simple, it includes one closing parenthesis. If it is complex, it includes more than one closing parenthesis.

First, it is a question of the rules related to the simple predication.

Distributional rule

This rule specifies the information which considers the predicate-argument structure by explicitly specifying the semantic nature of predicates and the different categories of arguments. There are three categories of predicates, monadic predicates, dyadic predicates and the triadic predicates, *cf. supra*³². It follows that there are three categories of arguments: the first argument which is coded X0 and it is observed with the three categories of

²⁹ WEATHER REPORT_FOG (DEIXIS+WEATHER REPORT).

³⁰ ABSENCE (HUMAN:DOCTOR).

³¹ ANNOUNCEMENT(HUMAN:DOCTOR,HUMAN:PATIENT,ABSENCE(HUMAN:DOCTOR)))

³² There is a fourth category of predicates, a predicate that its domain of arguments is quaternary. Its scarcity led to its neglecting [30].

predicates, the second argument, coded X1 and which is observed with the dyadic and triadic predicates, the third argument, coded X2 which is observed merely with one triadic predicate. The three predicate categories are respectively coded PRED(X0) monadic predicate, PRED(X0,X1) dyadic predicate and PRED (X0,X1,X2) triadic predicate. For example, the utterance *La météo est bonne* (*The weather is fine*) is generated from the following information: PRED=METEO%type_temps%(X0=METEO+X0=DEICTIQUE)³³.

The code %type_temps% means that the information is not in the linguistic database but depends on the knowledge base. For the X0 argument, it is about either a substantive of the METEO (WEATHER FORECAST) class or about a pronoun of the DEICTIQUE (DEICTIC) class. The exact nature of the deictic is specified at the level of the constructional rule (*règle constructionnelle*), cf. *infra*.

Pragmatic Adjustment rule

To limit the number of answers, the knowledge base is requested straightaway to identify a precised weather forecast type. As a result, the previous rule is rewritten as the following (in the case of a fine weather): PRED=BEAU_TEMPS(X0=METEO+X0=DEICTIQUE)³⁴.

Lexical rule

It is about the specification of the linguistic units associated with the semantic categories of predicates and arguments. For example, for the predicate class BEAU_TEMPS (FINE_WEATHER), the rule produces the following linguistic forms: the nouns *beau temps*, *temps magnifique* (*fine weather*, *a nice day*); the adjectives *agréable*, *beau*, *bon*, *magnifique* (*pleasant*, *delightful*, *nice*); or a verbal phrase *aller vers le beau* (*it is going to be beautiful*) and for the argumental class METEO (WEATHER FORECAST) the nouns *météo* and *temps* (*weather forecast* and *weather*).

Constructional rule

It is about the specification of the constructions connected to the linguistic forms of predicates. For example, in *beau temps* (*fine weather*), it is about the construction X0 :DEIXIS1 avoir DU N³⁵ as N is for *beau temps* (*fine weather*)³⁶, and the subject is the argument X0 of the nominal predicate under the value coded DEIXIS1, which means *on* or *nous* (*we*).

Reconstruction rule

The construction, associated with the linguistic form of a predicate, is canonical; it involves a standard layout of the subject position and, if any, the positions of the first and second complement occupied by the arguments. In speech, this layout may be changed and codified [31]. It is the case with the passive form in comparison to the active form of a sentence. For example, the predicate *petit déjeuner* (*breakfast*) is associated with the canonical construction: LE N être X0:GN³⁷ (*Le petit déjeuner est du café, des tartines et un jus d'orange*) (*The breakfast is*

³³ PRED=WEATHERFORECAST%type_weather%(X0=WEATHER FORECAST+X0=DEICTIC)

³⁴ PRED=FINE_WEATHER (X0=WEATHER FORECAST+X0=DEICTIC)

³⁵ X0: DEIXIS1 to have THE N

³⁶ *Avoir* (*to have*) is a helping verb of the nominal predicate [20].

³⁷THE N to be X0:NG

coffee, a toast and an orange juice) and with the reconstructions *il y avoir comme N X0:GN*³⁸ (*Il y a comme petit déjeuner du café, des tartines et un jus d'orange*) (*There is for breakfast: coffee, a toast and an orange juice*), etc.

The lexical variety of the predicates and arguments and the constructional variety, combined with each other, explain that from the same propositional content, the system produces a huge number of utterances and allows the stimulation of a human conversation, *cf. infra*.

Instantiation rule

This rule consists of merging the three previous rules and inserting the linguistic forms in the positions of the predicates and arguments. For example, in the previous example, the construction *LE N être X0: GN* is transformed into *LE petit déjeuner être DET café, DET tartines, DET jus d'orange*³⁹, the reconstruction *il y avoir comme N X0:GN* into *il y avoir comme petit déjeuner DET café, DET tartines, DET jus d'orange*⁴⁰, etc.

Actualization rule

It is about the specification of the predicate time and the argument determination. For example, when the instantiation rule gives such results : *aujourd'hui, DET météo être à la pluie/aujourd'hui, DET temps être à la pluie/DET météo être à la pluie aujourd'hui/DET temps être à la pluie aujourd'hui*⁴¹. The application of the code relative to the enunciative properties of the nominal predicate *pluie* (*rain*) produces the result : *aujourd'hui, LE météo être:PRESENT à la pluie/aujourd'hui, LE temps être:PRESENT à la pluie/LE météo être:PRESENT à la pluie aujourd'hui, LE temps être:PRESENT à la pluie aujourd'hui*⁴².

Conjugation and coordinating conjunctions rule

This rule specifies how to conjugate a verb, how to coordinate and use determiners.

Morphological adjustment rule

This rule specifies in which conditions, there are contractions, transformations or removal of concatenated forms and stop words (*mots vides*), for example:

Lexico-syntactic adjustment rule

The role of this rule is to limit the descriptive power of the syntactic properties, particularly the ones which concern the reconstructions and/or the lexical combinatorics. In fact, all the predicate uses defined by the same construction do not accept the reconstructions associated with them. Hence, it is convenient to limit them by only specifying the ones applicable. For example, adjective predicates *bon, mauvais, pourri* (*good, bad, rotten*) are characterized by the construction *X0: GN être A (X0: NG to be A)*. This allows the reconstruction *X0: GN être*

³⁸There to be as N X0:NG

³⁹THE breakfast to be DET coffee, DET toast, DET orange juice

⁴⁰There to have as breakfast DET coffee, DET toast, DET orange juice

⁴¹Today, DET weather forecast to be at the rain/today, DET weather to be at the rain/DET weather forecast to be at the rain today/DET weather to be at the rain today

⁴²Today, THE weather forecast to be: PRESENT at the rain/today, THE weather to be: PRESENT at the rain/ THE weather forecast to be: PRESENT at the rain today, THE weather to be: PRESENT at the rain today.

*très A and X0:GN être vraiment A (X0: NG to be very A and X0 : NG to be really A) in a way that the following utterances can be generated: La météo est bonne, La météo est très bonne, La météo est vraiment bonne, Le temps est bon, Le temps est très bon, Le temps est vraiment bon, La météo est mauvaise, La météo est très mauvaise, La météo est vraiment mauvaise, Le temps est mauvais, Le temps est très mauvais, Le temps est vraiment mauvais, La météo est pourrie, La météo est très pourrie, La météo est vraiment pourrie, Le temps est pourri, Le temps est très pourri, Le temps est vraiment pourri. (The weather forecast is good, The weather forecast is very good, The weather forecast is really good, the weather is good, the weather is very good, the weather is really good, the weather forecast is bad, the weather forecast is very bad, the weather forecast is really bad, the weather forecast is rotten, the weather forecast is very rotten, the weather forecast is really rotten). Yet, the utterances *Le temps est bon, Le temps est très bon, Le temps est vraiment bon* (the weather is good, the weather is very good, the weather is really good) are not acceptable (lexical combinatorics constraint) and *La météo est très pourrie* (the weather forecast is rotten) (reconstruction constraint). In addition, it is convenient to precise these constraints in a lexico-syntactic adjustment rule.*

In the case of a complex predication, the same rules are applied. However, once the rule of the predicate identification is applied, they have to process the *embedded predicate before the predicate which embeds a unit. When there are more than two predicates which are embedded, those rules must be applied from the most embedded predicate to the one whose level of embedding is the highest, cf. supra.*

5 Conclusion

The works that have been carried out as part of the chatbot development have contributed to set the foundations of a lexical generative grammar of the French language. The development of this grammar will allow the creation of other chatbot processing others themes.

References

1. Asimov, I.: Le cycle des robots. J'ai lu, Paris (2001).
2. Devillers, L.: Des robots et des hommes: mythes, fantasmes et réalité. Plon, Paris (2017).
3. Landragin, F.: Dialogue homme-machine conception et enjeux. Hermès science publications-Lavoisier, Paris (2013).
4. Kerbrat-Orecchioni, C.: Les interactions verbales. Armand Colin, Paris (1990).
5. Sacks, H., Schegloff, E.A., Jefferson, G.: A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*. 50, 696–735 (1974).
6. Sabah, G.: L'intelligence artificielle et le langage. Hermès, Paris (1989).
7. Pierrel, J.-M. ed: Ingénierie des langues. Hermès Science publications, Paris (2000).
8. Vilnat, A.: Quels processus pour les dialogues homme-machine. In: Sabah, G. (ed.) *Machine, langage et dialogue*. L'Harmattan, Paris (1997).
9. Henze, N., Dolog, P., Nejd, W.: Reasoning and ontologies for personalized e-learning in the semantic web. *Educational Technology & Society*. 7, 82–97 (2004).
10. Hoc, J.-M.: *Psychologie cognitive de la planification*. Presses universitaires de Grenoble, Grenoble (1987).
11. Benveniste, É.: *Problèmes de linguistique générale*. (1966).
12. Corblin, F.: *Les formes de reprise dans le discours: anaphores et chaînes de référence*. Presses Univ. de Rennes, Rennes (1995).
13. Grosz, B., Sidner, C.: *Attention, intentions, and the structure of discourse*, (1986).
14. Hiraoka, T., Neubig, G., Yoshino, K., Toda, T., Nakamura, S.: *Active Learning for Example-Based Dialog Systems*. In: Jokinen, K. and Wilcock, G. (eds.) *Dialogues with social robots*. pp. 67–78 (2017).

15. Chandramohan, S., Geist, M., Lefèvre, F., Pietquin, O.: User Simulation in Dialogue Systems Using Inverse Reinforcement Learning. (2011).
16. Gouritin, T.: L'arnaque chatbots durera-t-elle encore longtemps?, <https://www.frenchweb.fr/larnaque-chatbots-durera-t-elle-encore-longtemps/305697>, (2018).
17. Weizenbaum, J.: ELIZA - Un programme informatique pour l'étude de la communication en langage naturel entre l'homme et la machine. Communications de l'ACM (1966).
18. Bisk, Y., Yuret, D., Marcu, D.: Natural Language Communication with Robots. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. pp. 751–761. Association for Computational Linguistics, San Diego, California (2016).
19. Harris, Z.S.: Notes du cours de syntaxe. Seuil (1976).
20. Gross, M.: Les bases empiriques de la notion de prédicat sémantique. Langages. 15, 7–52 (1981).
21. Buvet, P.-A.: La dimension lexicale de la détermination en français. Honoré Champion, Paris (2013).
22. Mejri, S.: Le prédicat et les trois fonctions primaires. In: Souza Silva Costa, D. and Bençal, D.R. (eds.) Nos caminhos do léxico. Editora UFMS, Campo Grande do Sul (2016).
23. Martin, R.: Linguistique de l'universel: réflexions sur les universaux du langage, les concepts universels, la notion de langue universelle. Académie des inscriptions (2016).
24. Frege, G., Imbert, C.: Écrits logiques et philosophiques. Seuil, Paris (1971).
25. Fradin, B.: Nouvelles approches en morphologie. Presses Universitaires de France, Paris cedex 14 (2003).
26. Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal. 27, 379–423 (1948).
27. Buvet, P.-A.: Compréhension automatique des articles politiques : le traitement des discours rapportés. Ela. Études de linguistique appliquée. 180, 475–492 (2015).
28. Mel'čuk, I.: Cours de morphologie générale: théorique et descriptive. Presses de l'Université de Montréal, Montréal (1993).
29. Chomsky, N., Miller, G.A.: L'analyse formelle des langues naturelles. Mouton, Paris (1968).
30. Leclère, C.: Travaux récents en lexique-grammaire des verbes français. Duculot (1998).
31. Muller, C.: Le prédicat, entre (méta)catégorie et fonction. (2013).