

NATURAL LANGUAGE FOR FACTOID-WH IN ENGLISH AS A FOREIGN LANGUAGE

Ana L. Espinosa López

aespinosalopez@correo.unicordoba.edu.co

Adán A. Gómez Salgado

aagomez@correo.unicordoba.edu.co

Yina M. Vega Cálao

yinavegac@correo.unicordoba.edu.co

ABSTRACT

The purpose of this article is to describe the natural language made for the creation a Cognitive Model for Factoid-WH questions. A cognitive Model is a Representation of the cognitive processes that occur in the human mind. A Factoid-WH question is a question, which starts with WH-interrogated word and requires a fact as an answer. The methodology of this work is a cognitive modeling which consists of seven steps, three of them will be shown as the result of the analysis of the cognitive task in natural language for this model.

Keywords: Natural Language, Factoid-WH Question, EFL.

I. INTRODUCTION

A Factoid-Wh Question is a question, which starts with a Wh-interrogated word (What, When, Where, Who) and requires an answer as a fact expressed in the text body [1]. According to [2] Factoid-Wh questions have the following characteristics: i) Begin with an Interrogative Pronoun (IP) inverting the order between the subject and the operator and it is pronounced with falling intonation; ii) If there is not an auxiliary verb, “do or does” that acts as an operator is introduced; iii) “Be” and “Have” as lexical verbs which behaves the same as yes/no questions. Learners who are repeatedly exposed to Factoid-Wh questions generation during didactic activities of English as a Foreign language (EFL), produce new Factoid-Wh questions with a variety of lexical verbs and Wh-question words [3].

Question generation has been defined as task of automatically generating questions from some form of input [4]. This input is a written sequence,

resulting from a voice recognition system or obtained from a keyboard or even from a written document [5]. The input could vary from information in a database to a deep semantic representation to raw text. The Question Generation is viewed as a three-step process: content selection, question type selection and question construction [4]. Factoid-Wh Question Generation Process consists in receiving a text source as input, in order to automatically parsing the sentences and transforming these sentences into Factoid-Wh questions [6].

The problem identified for this research is centered in the difficulties that the second semester students of the “Licenciatura en Informática” at Universidad de Córdoba in Colombia have to formulate Factoid-Wh questions. This is evidenced by a diagnostic test that was applied to these students, where they have to construct a Factoid-Wh question in English as a Foreign Language. These test results show the difficulties that students present when constructing these types of questions which consist in the lack of knowledge of its basic structure, the inappropriate use of techniques to transform an affirmative affirmation into a Factoid-Wh question, and the confusions when differentiating grammatical structures

The development and implementation of this kind of cognitive educational resources based on models that represent cognitive processes have become the latest educational trend in Latin America [7]. This is because the principles and notions that are part of educational theories are used in the engineering process of this type of cognitive educational resources to design the reasoning and decision-making mechanisms on which these intelligent systems are based [8].

The purpose of this article is to present the cognitive task in natural language for the creation of the cognitive model for the Factoid-WH questions.

II. NATURAL LANGUAGE FOR FACTOID-WH QUESTIONS.

The methodology used is proposed by Olier, Gómez, & Caro (2018) [9] and consists of seven steps of which the first three phases are disclosed:

- **Selection Cognitive Task:** Obtaining Information for describing the Cognitive Task in Natural Language.
- **Obtaining Information for describing the Cognitive Task:** In this stage the information sources (obtained from experts, users or documental sources) in order to describe the cognitive task is selected.
- **Description of Cognitive Task in Natural Language:** This stage specifies in a natural way the necessary requirements in order to solve the problem.

Below is the development of the phases:

Cognitive Task Selected

The cognitive task to be modeled is the development of a cognitive model for the generation of Factoid- Wh question in English as a foreign language.

Information Obtained to detail the Cognitive Task

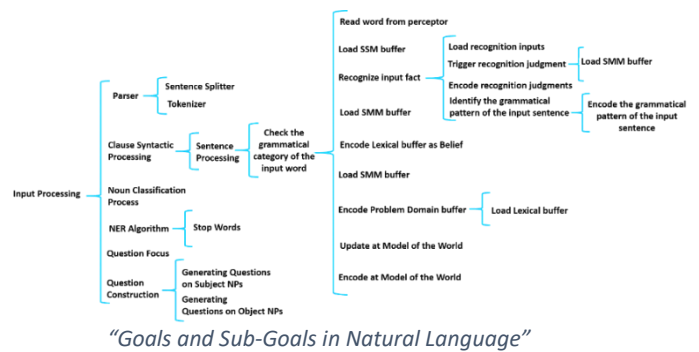
The information that describes the cognitive task was obtained through two experts and from some documentary sources.

Table 1. Format to synthesize the cognitive task description when the information source comes from experts.

Experts		X
Knowledge Area	Cognitive Computing and Applied Linguistics	
Number of Experts	1 MSc. in Technology of Information Applied to Education 2 Lic. in English	
Synthesis of Cognitive Task Description	Cognitive modeling is a research methodology of cognitive science, which produces theories that are expressed as computer programs. The central goals of cognitive modeling are: (a) describe (b) predict, (c) and prescribe human behavior through computational models of cognitive processes commonly called Cognitive Models.	

Description of Cognitive Task in Natural Language:

The cognitive model for Factoid-Wh questions is composed by the following elements: Goals, Actions, Mental States and Production Rules. This cognitive model has a main Goal called Input Processing which is structured by various sub-goals that allow to achieve the construction of Factoid-WH questions step by step. These Goals and sub-goals are presented below, considering the process of generating questions described by [10]: The Question Generation is viewed as a three-step process: Content Selection, Selection of Question type and Question Construction. In Figure. 1 is presented the goals and sub-goals that must be carried out for the construction of the questions Factoid-WH question.



Considering the main Goal and the Sub-goals presented, the cognitive model for the Factoid-WH question is described in natural language below.

1. The input is gotten and a Parsing process is developed:
 - a. The input is structured into sentences, thus word by word is read using a Parsing process called "Left Corner" and then the end of each sentence is detected by identifying a question mark, an exclamation point, "/ n", end point or the end of the text.
 - b. Words are separated from other textual elements such as parentheses, brackets, question marks, exclamation marks, numbers, digits, etc. Then, each of the remaining elements becomes Tokens. A token is each element that structures a sentence.
2. Each sentence is word by word syntactically processed, verifying the grammatical category of each detected word:
 - a. The word of the perceptor is read.
 - b. The word as an input fact is encoded.
 - c. The coded input fact is stored in the buffer of the sensory memory (SSM Buffer).
 - d. The input fact is copied from the SSM buffer to BCPU.Input
 - e. The input is loaded from the BCPU. Input (Buffer SSM)
 - f. A Pattern τ of the Pattern Set of the Short-Term Memory is loaded
 - g. The Belief β of the Semantic Memory (SMM) is retrieved using BCPU.Input as a cue.
 - h. The Belief β into the SMM Buffer is copied.
 - i. A recognition judgment is triggered ϑ only if the SMM Buffer is empty

- j. A new Belief β is encoded with the recognition judgment ϑ
 - k. The Belief ϑ into BCPU.Pattern is copied
3. The lexical buffer is loaded
 - a. The input data is loaded from the BCPU.Input (SSM)
 - b. The input fact is copied to <Word Node> Buffer / Field in MoW
4. The buffer of the problem domain is encoded.
5. The Belief β de Buffer / Campo in MoW is retrieved
6. The Belief β is copied in STM Lexical in MoW
7. The word node is updated in the MoW
 - a. The Belief β of <Word Node> Buffer/Field in MoW is retrieved
8. Word Node> is encoded in MoW
 - a. The Belief β of the Lexical Buffer is retrieved
 - b. The Belief β is loaded from Buffer/Field in MoW
 - c. The Belief β is copied in Lexical STM in MoW
 - d. The Belief β is copied to <Word Node> Buffer/Field in MoW
9. The classification of nouns is processed:
 - a. Search for nouns in the structure of the sentence.
 - b. Appropriate, not appropriate nouns and adverbs are labeled in the structure of the sentence.
10. Recognized Algorithm of Nominated Entities is executed
 - a. The output of the classification of nouns is obtained
 - b. Who / Whom is used for PERSON or ORGANIZATION or Where for LOCATION # If <Proper Noun> = true the algorithm is activated

- c. If the <Proper Noun> is false the algorithm is activated
 - d. The question word How is used is an adverb.
11. Connector words are chosen
- a. It is obtained from the output of the noun's classification process
 - b. The connecting words of the sentence are eliminated
 - c. Use of Who / Whom for PERSON or ORGANIZATION or Where for LOCATION or What in case THERE IS NO ENTITY
12. The question is focused
- a. The word of the question is identified
 - b. The question approach is selected
13. The question is generated
14. If the question Factoid WH-Question is subject NP is attached to the main verb of the sentence and is identified: NP = n1> (S = n2> ROOT) & \$ ++ VP = n3 (4)
15. If it is Factoid question WH-Question is NP objects are attached to the front of the sentence and the NP object is identified: NP = n1! >> NP >> (VP> (S = n2> ROOT))

III. CONCLUSION

In this paper, the cognitive process for the Factoid-WH questions generation when learning English as a foreign language is explained through a natural language.

Finally, with the development of this research, the modelling of the cognitive task in natural language for the creation of Factoid-WH questions generation process when learning English as a Foreign Language allows an advance from the cognitive informatics to education.

IV. REFERENCES

- [1] O. Kolomiyets and M.-F. Moens, "A survey on question answering technology from an information retrieval perspective," *Inf. Sci. (Ny)*, vol. 181, no. 24, pp. 5412–5434, 2011.
- [2] M. R. T. i Cherta, "Un estudio transversal de nivel de adquisición de la interrogación en inglés con alumnos de 8º de EGB," *Rev. española lingüística Apl.*, no. 9, pp. 169–187, 1993.
- [3] K. McDonough and J. De Vleeschauer, "prompt-type frequency, auditory pattern discrimination, and efl learners production of wh-questions," *Stud. Second Lang. Acquis.*, vol. 34, no. 3, pp. 355–377, 2012.
- [4] V. Rus and C. G. Arthur, "The question generation shared task and evaluation challenge," in *The University of Memphis. National Science Foundation*, 2009.
- [5] C. Ramos, J. C. Augusto, and D. Shapiro, "Ambient intelligence the next step for artificial intelligence," *IEEE Intell. Syst.*, vol. 23, no. 2, pp. 15–18, 2008.
- [6] Y. Skalban, L. Specia, R. Mitkov, and others, "Automatic question generation in multimedia-based learning," *Proc. COLING 2012 Posters*, pp. 1151–1160, 2012.
- [7] A. A. Gómez, M. F. Caro, A. M. Solano, and Y. M. Vega, "Trends of Educational Informatics in Latin America," *Int. J. Softw. Sci. Comput. Intell.*, vol. 10, no. 1, pp. 80–87, 2018.
- [8] A. A. Gómez and M. F. Caro, "Meta-Modeling Process of Pedagogical Strategies in Intelligent Tutoring Systems," in *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, 2018, pp. 485–494.
- [9] A. J. Olier, A. A. Gómez, and M. F. Caro, "Cognitive Modeling Process in Metacognitive Architecture CARINA," in *2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)*, 2018, pp. 579–585.

- [10] V. Rus, B. Wyse, P. Piwek, M. Lintean, S. Stoyanchev, and C. Moldovan, “A detailed account of the first question generation shared task evaluation challenge,” *Dialogue & Discourse*, vol. 3, no. 2, pp. 177–204, 2012.

