Article

Automatic Generation of Literary Sentences in French

Luis-Gil Moreno-Jiménez* [1] **0000-0001-7753-7349, Juan-Manuel Torres-Moreno***[1] **0000-0002-4392-1825, Roseli S. Wedemann**[2] **0000-0001-7532-3881**

[1] Laboratoire Informatique d'Avignon, Université d'Avignon, 339 Chemin des Meinajariès, 84911 Avignon, cédex 9, France; {luis-gil.moreno-jimenez, juan-manuel.torres}@univ-avignon.fr

[2] Instituto de Matemática e Estatística, Universidade do Estado do Rio de Janeiro, Rua São Francisco Xavier 524, 20550-900, Rio de Janeiro, RJ, Brazil; roseli@ime.uerj.br

* Correspondence: luis-gil.moreno-jimenez@univ-avignon.fr (L.G.M.J); juan-manuel.torres@univ-avignon.fr (J.M.T.M.)

Abstract: In this paper, we introduce a model for the automatic generation of literary sentences in French. It is based on algorithms that we have previously used to generate sentences in Spanish and Portuguese, and on a new corpus consisting of literary texts in French that we have constructed, called MEGALITE$^{\text{FR}}$. Our automatic text generation algorithm combines language models, shallow parsing and deep learning, artificial neural networks. We have also proposed and implemented a manual evaluation protocol to assess the quality of the artificial sentences generated by our algorithm, by testing if they fulfill four simple criteria. We have obtained encouraging results from the evaluators for most of the desired features of our artificially generated sentences.

Keywords: computational creativity; literary sentences; automatic text generation; shallow parsing and deep learning.

1. Introduction

Creativity is a phenomenon that has been addressed for the past few decades by the *Natural Language Processing* (NLP) scientific research community [1–3], and it is not easily defined. In *"The creative mind: Myths and Mechanisms"* [4], Maragaret Boden argues that the creative process is an intuitive path followed by humans to generate new artifacts that are valued for their novelty, significance to society and beauty. Although there have been major advances resulting from the research efforts that have been put forward to develop automatic procedures for producing creative objects, there are difficulties and limitations related to the inherent complexity of understanding the creative process in the human mind [5]. These difficulties hamper the task of modelling and reproducing the cognitive abilities employed in the creative process. The search for automated processes capable of creatively generating artifacts has recently given rise to a field of research called *Computational Creativity* (CC), which offers interesting perspectives in various fields of art such as the visual arts, music, and literature, among others.

The creation of literature is particularly interesting and difficult, when compared to other types of *Automatic Text Generation* (ATG) tasks, mainly due to the fact that literary texts are not perceived in the same way by different persons, and this perception can also vary depending on the reader's mood. Literary documents often refer to imaginary, allegorical or metaphorical worlds or situations, unlike journalistic or encyclopedic genres which mainly describe factual situations or events, and it is often difficult to establish a precise boundary between general language and literary language. It can thus be assumed that literary perception is subjective and it is therefore difficult to ensure that text generated by an algorithm will be perceived as literature. To reduce this possible ambiguity regarding literary perception, we consider that literature is regarded as text that employs a vocabulary which may vary largely when compared to that used in common language and that it employs various writing styles and figures of speech, such as rhymes, anaphora, metaphores, euphemism and many others, in order to obtain an artistic and possibly complex and emotional text [6]. This understanding has given us a guide for the development of our model whereby literary sentences can be generated.

Citation: Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S. Title. *Preprints* **2022**, *1*, 0. https://doi.org/

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

In this paper, we describe our model for the production of literary sentences based on homosyntax, where semantically different sentences present the same syntactic structure. This approach is different than paraphrasing and seeks to generate a new sentence which preserves a given syntactic structure by strongly modifying the semantics. Our proposal is based on a combination of language models, grammatical analysis and semantic analysis based on an implementation of *Word2vec* [7]. In Section 2, we present some current work regarding literary text production, with a focus on methods related to our own work. In Section 3, we present the corpus that we have constructed and used for training our algorithms. We explain the algorithms that we have developed for the production of literary sentences in Section 4. Some examples of the artificial sentences that are generated are shown in Section 5, along with the results obtained with evaluations done by humans. Finally, we present our conclusions and propose some ideas for future investigations in Section 6.

2. Related Work

In this section, we present some of the recent work with a focus on the production of literary and non-literary text. Early attempts to model the process of automatic text production were based on stochastic models, such as the one presented by Szymanski and Ciota [8], where text is generated in Polish from *n*-charactergrams combined with Markov chains. Other techniques have also been used for text generation, such as the one proposed in [9], by Molins and Lapalme, that is based on grammatical structures for text generation in French and English. This method is known as *canned text* and it provides a stable and consistent grammatical basis for text production. As the grammatical structure is previously determined, the efforts of the ATG algorithm is concentrated on selecting an appropriate vocabulary considering meaning.

More recently, many of the algorithms that have been proposed are based on neural networks. In [10], the algorithm that is proposed by Van de Cruys generates coherent text using a recurrent neural network (RNN) and a set of keywords. The keywords are used by the RNN to determine the context of the text that is produced. Another RNN approach, proposed by Clark, Ji and Smith in [11], is used to produce narrative texts, such as fiction or news articles. Here, the entities mentioned in the text are represented by vectors which are updated as the text is produced. These vectors represent different contexts and guide the RNN in determining which vocabulary to retrieve to produce a narrative. Fan, Lewis and Dauphin improve the efficiency of standard RNNs, by using a convolutional architecture, to build coherent and fluent passages of text about a topic in story generation [12].

Some very interesting proposals combine both the neural network based procedures and canned text methods [9]. This is also the case of Oliveira [13], who has made a thorough study of models for the automatic generation of poems, and proposed his own method of generation based on the use of canned text [14]. Another work based on canned text is presented in [15], for the production of stanzas of verse in Basque poetry. Finally, in [16], Zhang and Lapata propose an RNN for text structure learning for the production of Chinese poetry.

3. The MegaLite^{FR} Corpus

We now present the properties of the corpus that we have constructed for the training and validation of our literary ATG models. Although the constitution and the use of specifically literary corpora is very important for developing and evaluating algorithms for literary production, the need for such corpora has been systematically underestimated. Literary corpora are needed mainly due to the possible level of complexity of literary discourse, and the subjectivity and ambiguity aspects normally found in literary texts. The increased difficulty involved in producing these literary corpora usually induces users to resort to the use of corpora consisting of encyclopaedic, journalistic or technical documents for textual production. In order to have a substantial and appropriate resource for the

 doi:10.20944/preprints202212.0495.v1

generation of literary sentences in French, we have concentrated efforts in the construction of a corpus consisting solely of French literature called MEGALITE^{FR}.

Our MEGALITE^{FR} corpus consists of 2,690 literary documents in French, written by 620 authors [17]. Most of the documents were originally written in French, some were translated to French and an important part of this corpus comes from the Bibebook[1] ebook library. Some relevant properties of MEGALITE^{FR} are shown in Table 1, and the distribution of works by genre is shown in Table 2. These properties suggest that it is a corpus with an adequate size for training machine learning algorithms for ATG, and we thus consider that the MEGALITE^{FR} corpus, containing only literary documents, is well adapted to our purpose of producing literary sentences in French.

Table 1. Properties of MEGALITE^{FR}. M = 10^6 and K = 10^3.

	Documents	Phrases	Words	Characters	Authors
MEGALITE^{FR}	2,690	10 M	182 M	1081 M	620
Mean per document	-	3.6 K	67.9 K	401 K	

Table 2. Distribution of genres in MEGALITE^{FR}.

	Plays	Poems	Narratives
MEGALITE^{FR}	97 (3.61%)	55 (2.04%)	2,538 (94.35%)

4. A Literary ATG Model

We have developed algorithms for the production of literary sentences in Spanish and Portuguese [6,18–21] and, in this paper, we review some of their main features which we adapt here for the production of literary sentences in French. The algorithms of our ATG model use keywords (queries) provided by the user as a semantic guide that determines the semantic context of the phrases that are produced. The model involves two basic steps.

- The first step consists in generating Partially Empty Grammatical Structures (PGS) corresponding to the words of a chosen sentence. For this purpose, we have implemented a procedure based on the *canned text* method, which is efficient for parsing in ATG tasks [22].
- In the second step, each tag (morpho-syntactic label) of the PGS generated in the first step is substituted by an alternative word, which is selected by means of a semantic analysis carried out with the aid of a procedure based on Word2vec [7].

4.1. Canned Text Based Procedure

We prepared a set composed of sentences that are selected from MEGALITE^{FR} manually, respecting the following criteria.

- Each sentence must express a clear, concrete message, that does not need a prior context to be understood.
- Each sentence must have a length N, such that $5 \leq N \leq 10$.
- The sentence must contain at least 3 lexical words.

We choose a literary sentence in French, f, from this set. Sentence f is then parsed with FreeLing [23] in order to replace the lexical words [2] with their morpho-syntactic labels (POS tags) and, in this way, generate a PGS. The functional words[3] in f are not substituted and remain in the sentence that will be generated. Figure 1 illustrates this first step, where

[1] Site available under the Creatives Commons BY-SA license, http://www.bibebook.com
[2] Verbs, adjectives and nouns.
[3] Prepositions, pronouns, auxiliary verbs or conjunctions.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 December 2022 doi:10.20944/preprints202212.0495.v1

the filled boxes represent the functional words and the empty boxes represent the POS tags that replaced the lexical words.

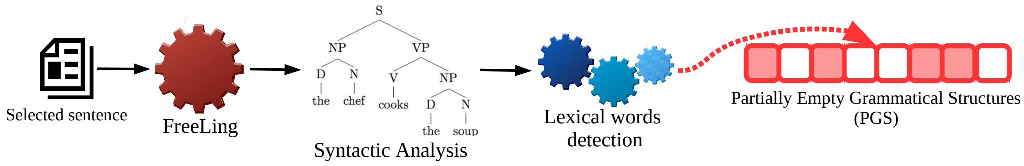

Figure 1. Illustration of the first step of our algorithm based on the *canned text* method.

4.2. Choice of Vocabulary to Generate the New Phrase

Once the PGS has been formed in the previous step, each of its POS tags are replaced by a word chosen from a vocabulary produced by an algorithm that implements a semantic analysis supported by the Word2vec method [7]. For the replacement, we used the method 3CosAdd, introduced by Drozd, Gladkova and Matsuoka in [24], that mathematically captures the analogy relations among words. This method considers the relationship between a set of words, for example, *France*, *Paris*, *Spain*, and a missing word x. Suppose that *France*, *Paris* and *Spain* are words that belong to the vocabulary of a corpus called **CorpA** that was used to train Word2vec, and consequently, $\overrightarrow{Paris}$, $\overrightarrow{France}$, and $\overrightarrow{Spain}$ are respectively the embedding vectors associated with these words after training. The word x is then determined by finding a vector $\vec{x}$ associated with a word in **CorpA**, such that $\vec{x}$ is closest to $\vec{y} = \overrightarrow{Paris} - \overrightarrow{France} + \overrightarrow{Spain}$, according to the cosine similarity between $\vec{y}$ and $\vec{x}$ given by

$$\cos(\vec{x}, \vec{y}) = \frac{\vec{x} \cdot \vec{y}}{||\vec{x}|| \; ||\vec{y}||} . \tag{1}$$

The answer to this specific example is considered correct if $\vec{x}$ corresponds to *Madrid* in the vocabulary of **CorpA**.

For our literary ATG algorithm, we will always refer to word embeddings that are produced by Word2vec when trained with MEGALITE$^{\text{FR}}$. Let us consider the words Q, o and A, where:

- Q represents the context specified by the user,
- o is the original word in f whose POS tag is to be replaced to generate the new phrase,
- A is the word previous to o in f.

These words are represented by the embeddings $\vec{Q}$, $\vec{o}$ and $\vec{A}$ that are used to calculate

$$\vec{y} = \vec{A} - \vec{o} + \vec{Q} . \tag{2}$$

Vector $\vec{y}$ has thus been produced by enhancing the features of $\vec{A}$ and $\vec{Q}$ and reducing the features of $\vec{o}$ ($\vec{y}$ is further away from $\vec{o}$). We then keep the first $M = 4,000$ embeddings that are closest to $\vec{y}$ in a list $\mathcal{L}$, i. e. we take the first M outputs of Word2vec when $\vec{y}$ is given as input. The list $\mathcal{L}$ thus has M entries, where each entry, $\vec{L_j}$, corresponds to the embedding of a word, w_j, associated to $\vec{y}$. The value of M was chosen as a compromise between the execution time and the quality of the results of the experiments that we have conducted. We next calculate the cosine similarity, θ_j, between each $\vec{L_j}$ and $\vec{y}$, according Equation (1) so that

$$\theta_j = \cos(\vec{L_j}, \vec{y}) . \tag{3}$$

List $\mathcal{L}$ is then ranked in descending order of θ_j.

If we replace the first POS tag in the PGS, A is an empty word, so $\vec{y} = -\vec{o} + \vec{Q}$. For example, for $Q = love$ and the sentence $f = I\ play\ the\ guitar$, when replacing the inflected verb $o = play$, we compute $\vec{y} = -\overrightarrow{play} + \overrightarrow{love}$ to obtain the ordered list $\mathcal{L}$. Some of the words corresponding to embeddings obtained in $\mathcal{L}$ for this example are *like*, *role*, *enchant*, *abandon*.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 December 2022 doi:10.20944/preprints202212.0495.v1

The words in this list are then joined with the words adjacent to o in the sentence f, and an algorithm that analyzes bigrams is used to choose which word will substitute o, as we describe next.

4.2.1. Bigram Analysis

An important feature to consider when choosing the word to replace the POS tag associated to word o is consistency. To reinforce consistency and coherence, we implemented a bigram analysis by estimating the conditional probability of the presence of the n^{th} word, w_n, in a sentence, given that another adjacent word, w_{n-1}, is present on the left of w_n, expressed as

$$P(w_n|w_{n-1}) = \frac{P(w_n \wedge w_{n-1})}{P(w_{n-1})}. \tag{4}$$

The conditional probability calculated by equation (4) corresponds to the frequency of occurrence of each bigram in MegaLite^{FR}. Among the bigrams of MegaLite^{FR}, we considered only those formed by lexical and functional words (punctuation, numbers and symbols were ignored) to create the list, LB, with all these bigrams, which we used to calculate the frequencies.

For each $\vec{L_j} \in \mathcal{L}$, we form two bigrams, $b1_j$ and $b2_j$, where $b1_j$ is formed by the word adjacent to the left of o in f concatenated with the word w_j. Then $b2_j$ is formed by concatenating w_j with the word adjacent to the right of o in f. To illustrate the formation of these bigrams, we show in the Figure 2 an example where f is the sentence in English *"I play the guitar"* and the word $o = play$ will be substituted. We then calculate the arithmetic mean, bm_j, of the frequencies of occurrence of $b1_j$ and $b2_j$ in LB. If o is the last word in f, bm_j is simply the frequency of $b1_j$. The value bm_j of each w_j is then combined with the cosine similarity θ_j, obtained with Equations (1) and (3), and the list $\mathcal{L}$ is reordered in descending order according to the new values

$$\theta_j = \frac{\theta_j + bm_j}{2}, \quad 1 \le j \le M. \tag{5}$$

We then take the first *embedding* in $\mathcal{L}$ that corresponds to the highest score, θ_{max}, and the best word, w_{max}, to replace o. Finally, a morphological analysis is performed with FreeLing, in order to transform the selected word, w_{max}, to the correct form according to the inflection specified by its corresponding POS tag (e.g., conjugations, gender, or number conversions).

With this procedure, we thus select the word that will substitute o that is semantically closest to $\vec{y}$, based on the analysis performed with Word2vec, while maintaining the coherence and consistency of the generated text by using the linguistic bigram analysis performed on the language model (the MegaLite^{FR} corpus). This process is repeated for each lexical word in f that is to be substituted, i.e., for each POS tag in the PGS. The result is a new sentence that does not exist in MegaLite^{FR}. The model is illustrated in Figure 3.

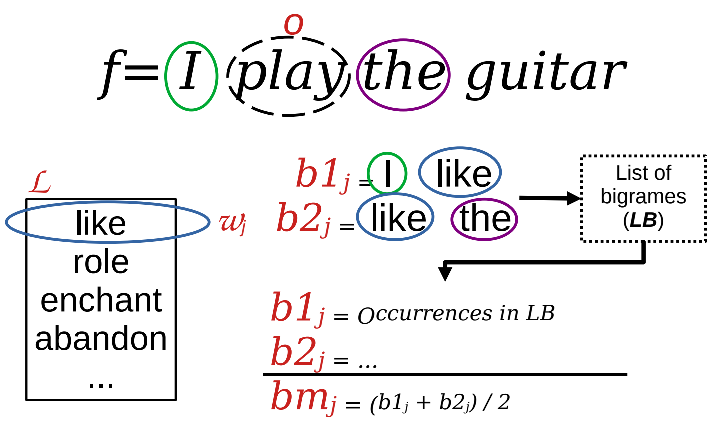

Figure 2. Illustration of the procedure based on bigram analysis for generating a new sentence.

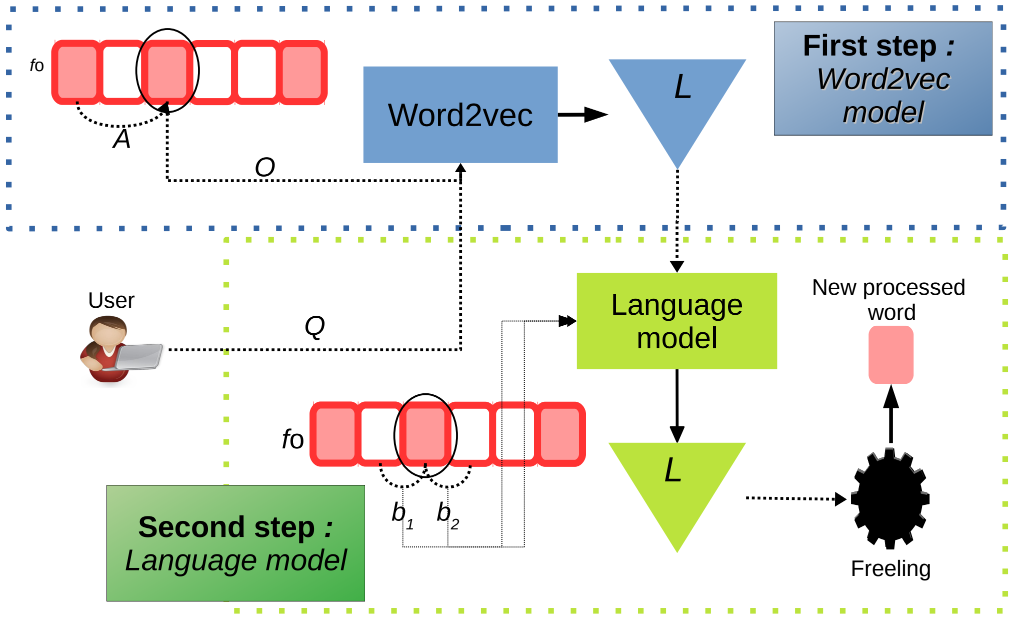

Figure 3. Illustration of the model for generating literary sentences.

5. Experiments

In this section, we present a manual protocol to evaluate the sentences produced by our model. Given the inherent subjectivity and ambiguity involved in literary perception, literary ATG algorithms developed for producing objects, in the domain of computational creativity, are frequently evaluated manually. We now show some examples and the results of the evaluation of sentences in French produced by our model, when using the MEGALITE^{FR} corpus for training the Word2vec method, the generation of the PGS's and the bigram analysis. The configuration of hyper-parameters for the Word2vec procedure used in these experiments is specified as follows:

- the number of learning periods executed with MEGALITE^{FR}, *iterations* = 10;
- the minimum number of times a word must appear in the corpus to be included in the model vocabulary, *minimum count* = 3;
- the dimension of the embedding vectors, *vector size* = 100;
- the radius of adjacent words that will be associated to a specific word in a sentence, during the training phase of the model, *window size* = 5.

Here are some examples of sentences in French, generated with three different PGS's and three queries, Q, that are shown in the format: sentence in French (*translation of sentence to English)*.

For Q = Love.

1. Il n'y a pas de passion sans impulsif. (*There is no passion without impulse.*).
2. Il n'y a pas d'affection sans estime. (*There is no affection without esteem.*)
3. Il n'y a ni aimante ni mélancolie en sérénité. (*There is neither love nor melancholy in serenity.*)
4. Il n'y a ni fraternelle ni inquiétude en anxiété. (*There is no brotherhood or concern in anxiety.*)

For Q = Sadness.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 December 2022 doi:10.20944/preprints202212.0495.v1

1. En solitude, la première tendresse est la plus forte. (*In solitude, the first tenderness is the strongest.*)
2. Il n'y a pas de confusion sans amour. (*There is no confusion without love.*)
3. Il n'y a pas de liaison sans ordre. (*There is no connection without order.*)
4. En douleur, la première mélancolie est la plus grande. (*In pain, the first melancholy is the greatest.*)

For $Q =$ Friendship.

1. En union, la première sollicitude est la plus belle. (*In union, the first concern is the most beautiful.*)
2. Il n'y a ni fraternelle ni faiblesse en impuissance. (*There is no brotherhood or weakness in powerlessness.*)
3. Il n'y a pas de sympathie sans émoi. (*There is no sympathy without emotion.*)
4. Il n'y a ni amie ni honte en peur. (*There is no friend or shame in fear.*)

In these examples, we can observe reasonably coherent sentences with words belonging to the same semantic field. There are some small syntax errors that occurred within the FreeLing tokenization module, which we expect may be solved by a fine-grained, a posteriori analysis, based on regular expressions.

5.1. Evaluation Protocol and Results

We have developed an evaluation protocol with four criteria that includes a Turing test. Evaluators were asked to evaluate the sentences considering: correct grammar, relation to context, literary perception and a Turing test. The set of sentences that were presented to the evaluators consisted of sentences generated by our literary ATG algorithm mixed with sentences generated by human writers, and this allowed us to compare our artificially generated sentences with non-artificial, human sentences.

We generated $p = 70$ artificial sentences for the evaluation set with our literary ATG model and with the following the contexts (queries): tristesse (sadness), amitié (friendship) and amour (love). We also asked 18 native French speakers, all of which have a university masters degree, to write sentences. These sentences written by humans were also based on the same previously generated PGS's that we used to generate the artificial sentences. So each person received six PGS's and was asked to write sentences by replacing the POS tags in the PGS's with French words, respecting the grammatical properties indicated by the POS tags. We thus obtained $p' = 230$ literary sentences written by human beings. We then randomly mixed the p artificial sentences with the p' manually created sentences and obtained a total of $P = 300$ sentences that constituted the evaluation set.

We then sent the sentences in the evaluation set to the evaluators (or annotators) many of which had also participated in the writing efforts, so we made sure that the annotators did not evaluate their own sentences. For the assessment of correct grammar and literary perception, the raters were asked to indicate if the sentences were bad or good (a binary response), according to their personal perception. The relationship to the context was assessed with three categories: poor relationship, good relationship, and very good relationship. Finally, for the Turing test the raters should indicate if they considered that the sentence was generated by a machine or not.

The results that we have received from the evaluators are encouraging. Approximately 80% of the artificial sentences were found to be grammatically correct, which is similar to the case of the sentences written by humans where 84% were classified as correct. The standard deviation calculated for the classification of the artificial sentences is 0.22 and indicates an acceptable level of agreement between evaluators. For the evaluation of literary perception, 44% of the artificial sentences were perceived as literary against 73% of the human sentences and, although the score obtained by the artificial sentences may seem low, it is still encouraging, when we consider the ambiguity inherent to literary perception among people (see Figure 4). However, this result indicates that there is still room for improvement regarding the literariness of the sentences generated by our ATG procedures.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 December 2022 doi:10.20944/preprints202212.0495.v1

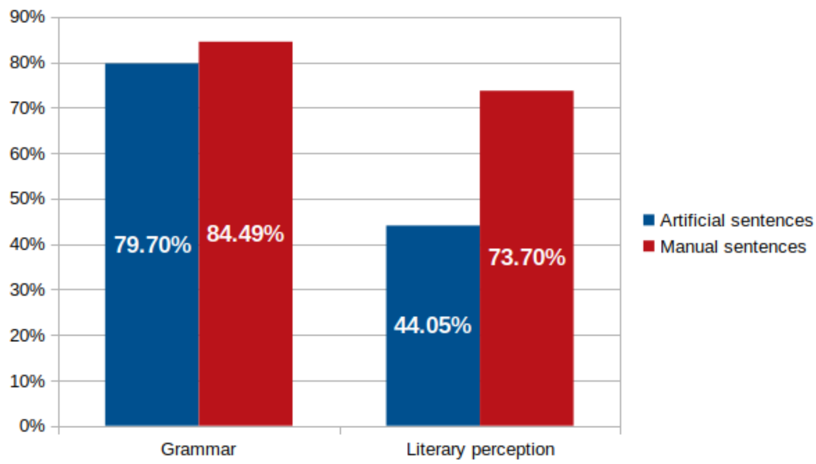

Figure 4. Results of the evaluation of grammatical correctness and literariness of the French sentences generated by humans and by our ATG model.

In Figure 5, we display the results obtained for the criterion that regards the semantic relationship of the artificial sentences with the contexts given by the user (the queries). It can be seen that the raters considered that 24% of the artificial sentences had a bad relation with the given context. On the other hand, the raters considered that 36% of the sentences have a good relation to context and almost 40% of the sentences were considered to have a very good relation to the expected context.

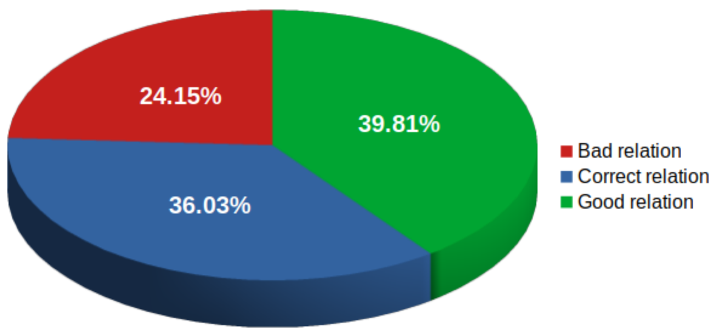

Figure 5. Results for the assessment of the semantic relation between the artificially generated sentences and the given query.

Finally, with respect to the Turing test, 44% of the artificial sentences were perceived as sentences generated by humans. Although at first it may seem otherwise, this result is also encouraging, when we consider that 26% of the human sentences were perceived as being artificial (see Figure 6). Given that only 74% of human sentences were perceived as sentences written by humans, the difference between the positive score of the ATG sentences and the human sentences is 30% and this gap suggests that possible improvements to our algorithm may produce similar perception of natural human sentences by evaluators for our ATG model.

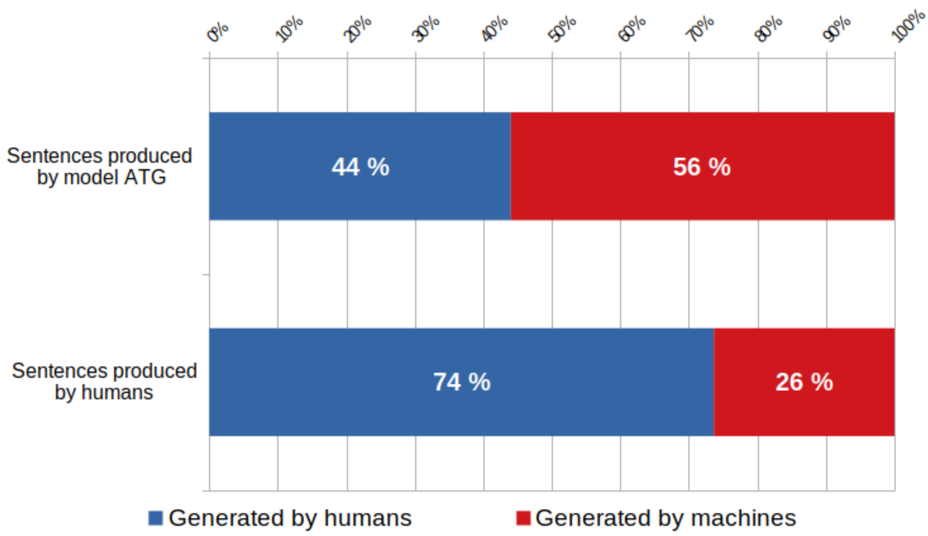

Figure 6. Results of the Turing test applied to the sentences produced by the ATG model. Red is used for the sentences perceived as artificial and blue for the sentences perceived as written by humans).

6. Conclusions

In the present work, we have described a model composed of algorithms that is capable of producing literary sentences in French, based on the canned text and Word2vec methods, and a bigram analysis of the MEGALITE^{FR} corpus that we have constructed. The model produces sentences with linguistic features frequently observed in literary texts written by humans.

Approximately 80% of the sentences generated by our model were perceived as being grammatically correct. In the case of the relationship to an expected context, approximately 76% of the sentences were classified as having a good contextual relationship. As for the Turing test, only 44% of the artificial sentences were perceived as having been written by a human. Although at first this seems not to be a good result, it is necessary to take into account the difficulty of this test, considering that the evaluation is also subjective. According to the literature, misleading humans in 50% of the evaluated texts is considered an acceptable result [25]. If we note that in our tests only 74% of manually produced sentences were perceived as having been written by a human being, the percentage of artificial sentences generated by our model that were considered to have been written by a human is 30% lower, and this can be considered as a margin of improvement for our algorithms in our future developments. It is also possible to invision plans to produce other literary structures such as rhymes and paragraphs with these types of algorithms.

Author Contributions: Conceptualization, Luis-Gil Moreno-Jiménez and Juan-Manuel Torres-Moreno; Data curation, Juan-Manuel Torres-Moreno; Formal analysis, Roseli Wedemann; Funding acquisition, Juan-Manuel Torres-Moreno; Investigation, Luis-Gil Moreno-Jiménez and Juan-Manuel Torres-Moreno; Methodology, Luis-Gil Moreno-Jiménez, Juan-Manuel Torres-Moreno and Roseli Wedemann; Project administration, Juan-Manuel Torres-Moreno; Resources, Juan-Manuel Torres-Moreno; Software, Luis-Gil Moreno-Jiménez and Juan-Manuel Torres-Moreno; Supervision, Juan-Manuel Torres-Moreno and Roseli Wedemann; Validation, Luis-Gil Moreno-Jiménez, Juan-Manuel Torres-Moreno and Roseli Wedemann; Writing – original draft, Luis-Gil Moreno-Jiménez, Juan-Manuel Torres-Moreno and Roseli Wedemann; Writing – review editing, Luis-Gil Moreno-Jiménez, Juan-Manuel Torres-Moreno and Roseli Wedemann.

Funding: This work is funded by the Consejo Nacional de Ciencia y Tecnología (Conacyt, Mexico), grant number 661101, and partially funded by the Université d'Avignon, Laboratoire Informatique d'Avignon (LIA), Agorantic programme (France). It has also received partial financial support from the Brazilian science funding agencies: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES).

Data Availability Statement: The MegaLite^{FR} corpus can be freely downloaded from the Ortolang webstie https://www.ortolang.fr/market/corpora/megalite#!. The versions in Spanish and Portuguese are also available.

Acknowledgments: R.S.W is also grateful for the kind hospitality of the Laboratoire Informatique d'Avignon (LIA), Université d'Avignon, where she participated in the conduction of this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sharples, M. *How We Write: Writing as creative design*; Routledge: London, UK, 1996.
2. Sridhara, G.; Hill, E.; Muppaneni, D.; Pollock, L.; Vijay-Shanker, K. Towards automatically generating summary comments for Java methods. In Proceedings of the IEEE/ACM International Conference on Automated Software Engineering; Association for Computing Machinery: New York, NY, USA, 2010; pp. 43–52. https://doi.org/10.1145/1858996.1859006.
3. Mikolov, T.; Zweig, G. Context dependent recurrent neural network language model. In Proceedings of the 2012 IEEE Spoken Language Technology Workshop (SLT); IEEE: New York, NY, USA, 2012; pp. 234–239. https://doi.org/10.1109/SLT.2012.6424228.
4. Boden, M.A. *The creative mind: Myths and Mechanisms*; Routledge, 2004.
5. de Carvalho, L.A.V.; Mendes, D.Q.; Wedemann, R.S. Creativity and Delusions: The dopaminergic modulation of cortical maps. In Proceedings of the 2003 International Conference on Computational Science (ICCS 2003).; Sloot, P.M.A.; Abramson, D.; Bogdanov, A.V.; Dongarra, J.J.; Zomaya, A.Y.; Gorbachev, Y.E., Eds.; Springer: Berlin, Heidelberg, 2003; Vol. 2657, pp. 511–520. https://doi.org/10.1007/3-540-44860-8_53.
6. Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S.; SanJuan, E. Generación automática de frases literarias. *Linguamática* **2020**, *12*, 15–30. https://doi.org/10.21814/lm.12.1.308.
7. Mikolov, T.; Yih, W.t.; Zweig, G. Linguistic regularities in continuous space word representations. In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL – HLT 2013); Association for Computational Linguistics: Atlanta, GA, USA, 2013; pp. 746–751.
8. Szymanski, G.; Ciota, Z. Hidden Markov models suitable for text generation. In Proceedings of the WSEAS International Conference on Signal, Speech and Image Processing (WSEAS ICOSSIP 2002); Mastorakis, N.; Kluev, V.; Koruga, D., Eds.; WSEAS Press: Athens, Greece, 2002; pp. 3081–3084.
9. Molins, P.; Lapalme, G. JSrealB: A bilingual text realizer for Web programming. In Proceedings of the 15th European Workshop on Natural Language Generation (ENLG); Association for Computational Linguistics: Brighton, UK, 2015; pp. 109–111. https://doi.org/10.18653/v1/W15-4719.
10. Van de Cruys, T. Automatic poetry generation from prosaic text. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Online, 2020; pp. 2471–2480. https://doi.org/10.18653/v1/2020.acl-main.223.
11. Clark, E.; Ji, Y.; Smith, N.A. Neural text generation in stories using entity representations as context. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL – HLT 2018); Association for Computational Linguistics: New Orleans, LA, USA, 2018; Vol. 1, pp. 2250–2260. https://doi.org/10.18653/v1/N18-1204.
12. Fan, A.; Lewis, M.; Dauphin, Y. Hierarchical neural story generation. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics; Association for Computational Linguistics: Melbourne, Australia, 2018; Vol. 1, pp. 889–898. https://doi.org/10.18653/v1/P18-1082.
13. Oliveira, H.G. A survey on intelligent poetry generation: Languages, features, techniques, reutilisation and evaluation. In Proceedings of the 10th International Conference on Natural Language Generation (CNLG); Association for Computational Linguistics: Santiago de Compostela, Spain, 2017; pp. 11–20. https://doi.org/10.18653/v1/W17-3502.
14. Oliveira, H.G.; Cardoso, A., Poetry generation with PoeTryMe. In *Computational Creativity Research: Towards Creative Machines*; Besold, T.R.; Schorlemmer, M.; Smaill, A., Eds.; Atlantis Press: Paris, France, 2015; Vol. 7, pp. 243–266. https://doi.org/10.2991/978-94-6239-085-0_12.
15. Agirrezabal, M.; Arrieta, B.; Astigarraga, A.; Hulden, M. POS-tag based poetry generation with WordNet. In Proceedings of the 14th European Workshop on Natural Language Generation (ENLG); Association for Computational Linguistics: Sofia, Bulgaria, 2013; pp. 162–166.
16. Zhang, X.; Lapata, M. Chinese poetry generation with recurrent neural networks. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP); Association for Computational Linguistics: Doha, Qatar, 2014; pp. 670–680. https://doi.org/10.3115/v1/D14-1074.
17. Moreno-Jiménez, L.G.; Torres-Moreno, J.M. MegaLite-2: An extended bilingual comparative literary corpus. In *Intelligent Computing*; Arai, K., Ed.; Springer: Cham, Switzerland, 2022; pp. 1014–1029. https://doi.org/10.1007/978-3-030-80119-9_67.

 doi:10.20944/preprints202212.0495.v1

18. Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S. Literary natural language generation with psychological traits. In *Natural Language Processing and Information Systems - NLDB 2020. LNCS*; Métais, E.; Meziane, F.; Horacek, H.; Cimiano, P., Eds.; Springer: Cham, 2020; Vol. 12089, pp. 193 – 204. https://doi.org/10.1007/978-3-030-51310-8_18.
19. Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S. Generación automática de frases literarias: un experimento preliminar. *Procesamiento del Lenguaje Natural* **2020**, *65*, 29–36.
20. Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S. A preliminary study for literary rhyme generation based on neuronal representation, semantics and shallow parsing. In Proceedings of the XIII Brazilian Symposium in Information and Human Language Technology and Collocated Events (STIL 2021); Ruiz, E.E.S.; Torrent, T.T., Eds.; Sociedade Brasileira de Computação: Online, 2021; pp. 190 – 198. https://doi.org/10.48550/arXiv.2112.13241.
21. Morgado, I.; Moreno-Jiménez, L.G.; Torres-Moreno, J.M.; Wedemann, R.S. MegaLite$^{\mathrm{PT}}$: A corpus of literature in Portuguese for NLP. In Proceedings of the 11th Brazilian Conference on Intelligent Systems, Part II (BRACIS 2022), LNAI; Xavier-Junior, J.C.; Rios, R.A., Eds.; Springer International Publishing: Cham, Switzerland, 2022; Vol. 13654, pp. 251–265. https://doi.org/10.1007/978-3-031-21689-3_19.
22. van Deemter, K.; Theune, M.; Krahmer, E. Real versus template-based natural language generation: A false opposition? *Computational Linguistics* **2005**, *31*, 15–24. https://doi.org/10.1162/0891201053630291.
23. Padró, L.; Stanilovsky, E. FreeLing 3.0: Towards wider multilinguality. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12); European Language Resources Association (ELRA): Istanbul, Turkey, 2012; pp. 2473–2479.
24. Drozd, A.; Gladkova, A.; Matsuoka, S. Word embeddings, analogies, and machine learning: Beyond king - man + woman = queen. In Proceedings of the 26th International Conference on Computational Linguistics: Technical Papers (COLING 2016); The COLING 2016 Organizing Committee: Osaka, Japan, 2016; pp. 3519–3530.
25. Turing, A.M. Computing machinery and intelligence. In *Parsing the Turing test*; Springer, 2009; pp. 23–65.