

Communication Mediated through Natural Language Generation in Big Data Environments: The Case of Nomao

Jean-Sébastien Vayre¹, Estelle Delpech², Aude Dufresne³, Céline Lemerrier⁴

¹CERTOP (UMR-5044), University of Toulouse Jean Jaurès, Toulouse, France

²Human Factors, Airbus, Blagnac, France

³LRCM, University of Montréal, Montréal, Québec

⁴CLLE (UMR-5263), University of Toulouse Jean Jaurès, Toulouse, France

Email: jean-sebastien.vayre@univ-tlse2.fr, delpech.estelle@gmail.com,
aude.dufresne@umontreal.ca, celine.lermercier@univ-tlse2.fr

How to cite this paper: Vayre, J.-S., Delpech, E., Dufresne, A. and Lemerrier, C. (2017) Communication Mediated through Natural Language Generation in Big Data Environments: The Case of Nomao. *Journal of Computer and Communications*, 5, 125-148.

<https://doi.org/10.4236/jcc.2017.56008>

Received: February 3, 2017

Accepted: April 27, 2017

Published: April 30, 2017

Copyright © 2017 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Along with the development of big data, various Natural Language Generation systems (NLGs) have recently been developed by different companies. The aim of this paper is to propose a better understanding of how these systems are designed and used. We propose to study in details one of them which is the NLGs developed by the company Nomao. First, we show the development of this NLGs underlies strong economic stakes since the business model of Nomao partly depends on it. Then, thanks to an eye movement analysis conducted with 28 participants, we show that the texts generated by Nomao's NLGs contain syntactic and semantic structures that are easy to read but lack socio-semantic coherence which would improve their understanding. From a scientific perspective, our research results highlight the importance of socio-semantic coherence in text-based communication produced by NLGs.

Keywords

Big Data, Natural Language Generation, Socio-Semantic Coherence, Cognitive Load, Reading, Eye Tracking

1. Introduction

Capitalism as we know it today is often referred to as cognitive based [1]. Thus knowledge about information management and mass communication brings considerable advantages to the actors in different spheres of our society. Consequently, Natural Language Generation (NLG) is a hot topic in today's big data movement. Indeed, NLG is a way to facilitate access to big data by transforming

it into a human-readable and semantically adequate form. Once in the form of a text, big data have an improved socio-economic value.

For instance, NLG is increasingly used in the field of data-journalism [2], *i.e.* the case of “robot journalists” like Quakebot for the Los Angeles Times, Wordsmith for the Associated Press, Quill for the US business magazine Forbes or Data 2 Content for the French newspaper Le Monde. In addition, some companies like Yseop in France offer decision support systems that are based on NLG and automatically generate customized reports for their customers.

Thus, in the light of these recent developments in the field of mass communication, it seemed appropriate to research how texts can be automatically generated and how these texts are read and understood by end-users. To serve this goal, we propose in this paper to study the case of the Natural Language Generation system (NLGs) developed by the company Nomao.

Nomao has edited since 2007 an online directory that lists local businesses in France. This directory exists in the form of a mobile and a web application. Nomao’s NLGs plays a major role in the production of the text content of this directory. Indeed, the role of the NLGs is to transform the raw data collected by Nomao about businesses into human readable texts. This paper addresses several innovative issues in the field of humanities: how does Nomao’s NLGs work? What are the socio-economic stakes that have guided its development? How do end-users read and understand the texts generated by the system?

In the first section, we describe the functioning of Nomao’s NLGs and identify related economic stakes. We show that Nomao’s NLGs is powerful. Yet, it has never been evaluated from a user perspective (Section 1). Therefore, we propose in Section 2 a theoretical and methodological framework for evaluating the quality of Nomao’s texts from a user perspective. In Section 3, we describe the experimental setup, which was deployed to conduct this evaluation. In Section 4, we show that while Nomao’s texts contain easy-to-read syntactic and semantic structures, they still lack socio-semantic coherence which could improve their understanding. We conclude this paper by underlying the fact that, from a scientific perspective, our research results underline the importance of socio-semantic coherence in text-based communication produced by NLGs.

2. Research Issues

As stated above, Nomao’s NLGs is of primary importance in the functioning of Nomao’s application. Nomao’s NLGs selects, combines and regenerates in the form of a text the data about local businesses that have been previously collected by Nomao. Thus, Nomao’s NLGs generates a large part of the textual content that is present in Nomao’s directory. Therefore, Nomao’s NLGs plays a major role between users and business owners since it implements and release, through a short descriptive text, the representation of the businesses listed in Nomao’s directory (cf. **Table 2**; Text 1).

To better understand the socio-cognitive and technical issue that underlies the evaluation of Nomao’s NLGs, we propose to describe its functioning and the

business model that influenced its design. Then we present the performance indicators that were used to evaluate the system so far and we will see that these indicators say nothing, from the user perspective, about the quality of the texts produced by this NLGs.

2.1. Nomao’s Natural Language Generation System (NLGs)

As shown in **Figure 1**, Nomao’s NLGs has been designed in accordance with the typical stages of NLGs [3]. These stages are: macro-planning (content selection followed by document structuring); micro-planning (syntactic planning, lexicalization, aggregation, referring expression generation) and surface realization (morphological adaptation, formatting).

Macro-planning is carried out in two sub-steps. First, content selection consists in selecting pieces of information that are to be conveyed to the end-user and which are relevant with regard to the communicative goals of the text. Second, document structuring consists in organizing the informational entities and establishing the rhetorical structure of the text. This first stage deals with the content of the text.

The goal of micro-planning is to determine how the informational entities selected and organized during macro-planning stage will be expressed in natural language. This second stage is carried out in four sub-steps: syntactic planning, lexicalization, aggregation and the generation of referential expressions. The second stage deals with the form of text.

Surface realization consists in operations that will transform the raw text outputted by microplanning into the final text. This stage includes several sub-steps: morphological adjustment (eg, generation of inflected forms through gender/number or verb/subject agreements), typographical adjustment (spaces, punctuation) and formatting (bold, caps, underline). This last stage provides the final form of the text.

2.2. Macro-Planning

The first sub-step of macro-planning retrieves pieces of information about the businesses that are listed in Nomao’s directory. To do this, the NLGs queries dedicated databases.

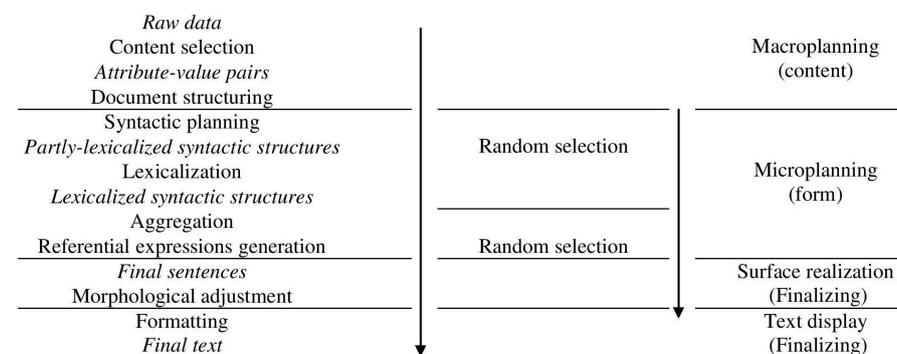


Figure 1. Nomao’s NLGs processing.

More precisely, the NLGs was set up by its designers so that it selects the following pieces of information: business identifier (e.g., “22370”), name (e.g., “Chai Saint Sauveur”), street name/number (e.g., “30 rue Bernard Mule”), city (e.g., “Toulouse”), nearby metro stations (e.g., “François Verdier”), nearby businesses (e.g., “category: eating, distance: 50m, name: Afro Resto”; “category: eating, distance 100m, name: Patricia Général Alimentation”), category (e.g., “eating <restaurant < French cuisine < south west”), tags (e.g., “eating, average price, reservation, with terrace, wine bar”) and terms (e.g., “terrace, bar, beautiful decor, cb visa mastercard”). These data items correspond to the informational entities that will eventually constitute the content that will be delivered to the end-users.

The second sub-step of macro-planning consists in building the rhetorical structure of the text. Nomao NLGs’ designers have identified five types of paragraphs that structure the final text. These paragraphs are:

- the title paragraph which is composed of the name of the business and the name of the city where it is located;
- the introductory paragraph that contains informational units describing the type of business and its location;
- the main paragraph which provides the user with various information about the business (e.g., the atmosphere, accepted payment methods, served dishes and drinks, etc.);
- a paragraph that describes subway stations in the vicinity of the business;
- an “opinion” paragraph that contains informational units about the business e-reputation.

2.3. Micro-Planning

As exposed above, the micro-planning stage begins with a syntactic planning sub-step. The NLGs selects a syntactic pattern for each of the informational units selected during macro-planning. This selection is made randomly among a set of patterns that are predefined and correspond to a specific informational entity. As shown in the example below, each pattern contains information about the syntactic relations underlying the organization of the various elements composing each statement.

It is worth noticing that some of the syntactic patterns can be nested into each other. For example, the pattern dedicated to the payment methods (cf. **Figure 2**) will be merged with the “ACCEPTED_PAYMENTS” pattern of **Figure 3**.

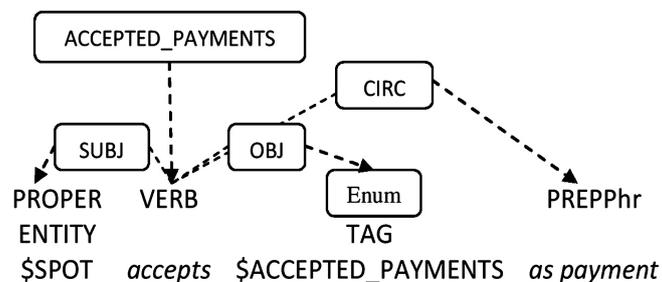


Figure 2. Example 1 syntactic pattern.

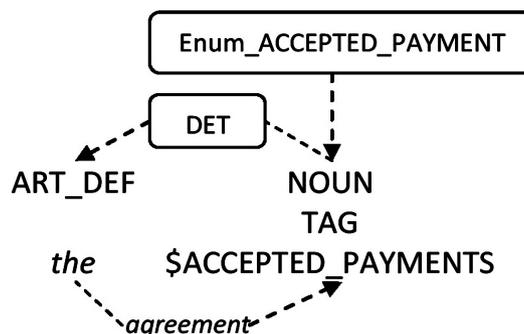


Figure 3. Example 2 syntactic pattern.

The lexicalization sub-step consists in inserting values in place of the variables contained in the syntactic patterns selected during the syntactic planning step. For example, the variables called “\$SPOT” and “\$ACCEPTED_PAYMENT” (cf. Figure 2) will be respectively replaced by their corresponding values: “restaurant” or “Chai Saint Sauveur” will replace \$SPOT and “Credit Card” or “VISA” will replace \$ACCEPTED_PAYMENTS.

In addition, morphosyntactic labels like “VERB” and “PREPPhr” will be replaced by corresponding natural language expressions randomly selected from a set of synonyms. The sets of synonyms have been previously defined by an expert linguist during the development of the NLGs. The synonyms were manually collected using the dictionary of the Research Center on Inter-languages Meaning in Context (CRISCO). For example, the verb “enjoy” in the statement “you will enjoy the draught beers of this bar” may be replaced by verbs like “appreciate” or “love”.

Each of the values that replace the variables come with morphosyntactic features specifying gender and number (eg, “restaurant, masc: sing,” “atmosphere, fem: sing”, etc.) in order to carry out grammatical agreement during the morphological adjustment step. As shown below, these morphosyntactic features will be used to generate the inflected form of the articles that agree with these nouns.

As illustrated in Figure 4, the goal of the aggregation step is to avoid generating a cumbersome text by merging repetitive structures (e.g., sentences with identical subjects and verbs).

The referential expressions generation step has a similar goal: it avoids cumbersomeness and redundancy by eliminating the repetition of identical subjects. It replaces repetitive noun phrases by corresponding personal pronouns or referential expressions. For example, the sequence of sentences “Chai Saint Sauveur is a restaurant [...]. Chai Saint Sauveur specializes in [...]. Chai Saint Sauveur offers dishes [...]” becomes: “Chai Saint Sauveur is a restaurant [...]. It specializes in [...]. This restaurant offers [...]”

The referential expressions that replace the initial noun phrase (“Chai Saint-Sauveur”) are chosen at random in a set of referential expressions that are semantically adequate to replace the initial noun phrase (e.g., “this restaurant”, “this business”, “this establishment”).

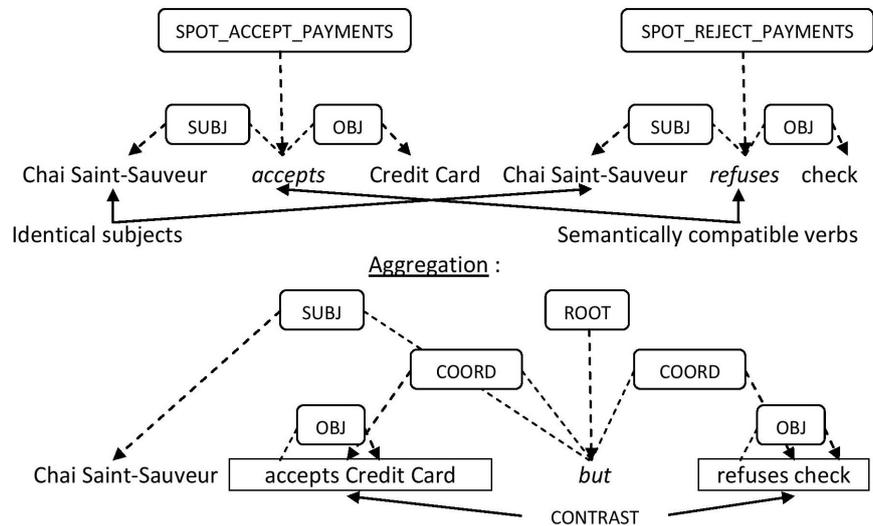


Figure 4. Aggregation example.

2.4. Surface Realization

The output of the micro-planning stage is a sequence of lemmas (non-inflected words) which should be morphologically adjusted.

Thus, the role of the surface realization stage is to apply agreement rules and perform surface operations like elisions and crasis so as to generate final word forms. For example, the French sentence “ce bar vous propose de faire une pause autour de un bon bière” becomes “ce bar vous propose de faire une pause autour d’une bonne bière” (where “un” has become “une”—article-noun agreement—and “de” has become “d”—elision).

Typographical adjustments apply classical typographic rules like uppercasing the first letter of each sentence and adding spaces between words. The final text is then ready to be displayed to end-users in the form of a short written description (cf. Table 3; Text 1) located in the web page of the corresponding business in Nomao’s online directory.

2.5. Design and Evaluation of Nomao’s NLGs

Nomao developed these NLGs in order to ensure the sustainability of its business model which consists in putting end-users in contact with businesses through premium rate telephone numbers. The company has not built any user community. Typical end-users land on Nomao’s website after having entered a local search query on a general-purpose web search engine (e.g., “restaurant in toulouse” “restaurant Chai saint sauveur toulouse”). It is thus obvious that the content of Nomao’s online directory needs to be optimized so that search engines will reference it well and suggest Nomao’s directory at the top of their search results. Good Search Engine Optimization (SEO) largely depends on the size and on the quality of the textual content that is inside a webpage: the better the content, the higher the ranking in the Search Engine Results Page (SERP).

Referring to the work of Cardon [4], NLG is an interesting manner to efficiently produce textual content and thus obtain a better control of SERP ranking.

However, this content must contain variations in expression for at least two reasons: first, end-users should not perceive that the contents of Nomao's directory have been machine-generated; otherwise this might hinder its acceptability by them. Second, Nomao's directory should not be regarded as malicious by search engines and be "blacklisted", especially by major search engines for which content quality is of high importance.

Thus, random lexical choices that are part of Nomao's NLGs ensure a certain variety of content as we can note with the two business descriptions below:

- V1 translation: "The best information about the "Mulligans" pub in Toulouse... Have a break in the pub "Mulligans" located 42 Rue Des Saules in the pleasant city of Toulouse. **This establishment proposes various beers.** The nearest metro station to this pub is Saint-Michel Marcel Langer. Near this pub you will find other places like De Danu or the Rosanna bar." (In french: "Le meilleur des infos sur le pub "Mulligans" à Toulouse... Détendez-vous le temps d'une pause dans le pub "Mulligans" localisé 42 Rue Des Saules dans l'agréable ville de Toulouse. **Différentes variétés de bières vous seront proposées dans cet établissement.** La station de métro la plus près de ce pub est Métro Saint-Michel Marcel Langer. Dans les environs de ce pub vous pourrez trouver le lieu de divertissement De Danu ou encore le bar Rosanna.")
- V2 translation: "Do you know the pub "Mulligans" in Toulouse? Relax at the pub "Mulligans" situated at 42 Rue Des Saules in the beautiful city of Toulouse. **The beer menu offers different varieties of beers to enjoy alone or with friends.** The nearest subway station is Metro Saint-Michel Marcel Langer. Feel free to check for other places nearby, like the De Danu pub or Rosanna bar." (In french: "Connaissez-vous le pub "Mulligans" à Toulouse ? Détendez- vous le temps d'une pause dans le pub "Mulligans" localisé 42 Rue Des Saules dans la jolie ville de Toulouse. **La carte de ce pub propose différentes variétés de bières à apprécier en solitaire ou entre amis.** L'arrêt de métro le plus près de ce pub est Métro Saint-Michel Marcel Langer. N'hésitez pas à repérer les autres possibilités aux alentours. Le lieu de divertissement De Danu ou le bar Rosanna se trouvent près de ce pub.")

Text variety is quantitatively monitored by Nomao with the Dice coefficient which measures similarity between two text samples. This coefficient is between 0 and 1. The closest to 1, the highest the similarity between the two text samples. More specifically, the Dice coefficient between a text sample X and a text sample Y is defined as twice the intersection between x and y divided by the sum of x and y; and where x is the set of open-class words uni-bi- and tri-grams in text sample X and y is the set of open word uni-bi- and tri-grams in text sample Y. An open-class word is either a verb, a participle, a noun, an adjective or an adjective-derived adverb. An n-gram is a sequence of n words. Dice coefficient formula:

$$s = 2|x \cap y| / (|x| + |y|) \quad (1)$$

For example, the texts X = "The restaurant is located in the beautiful city of

Toulouse” and $Y = \text{“This restaurant is in the city of Toulouse”}$ are respectively composed of 5 {restaurant, located, beautiful, city, Toulouse} and 3 {restaurant, city, Toulouse} open-class words. Their respective uni-bi- and tri-grams sets are: $x = \{\text{restaurant, located, beautiful, city, Toulouse, restaurant located, located beautiful, beautiful city, city Toulouse, restaurant located beautiful, located beautiful city, beautiful city Toulouse}\}$ and $y = \{\text{restaurant, city, Toulouse, restaurant city, city Toulouse, restaurant city Toulouse}\}$. The intersection of the sets x and y is therefore the set $x \text{ inter } y = \{\text{restaurant, city, Toulouse, city Toulouse}\}$. Therefore, the Dice coefficient corresponding to this example is as follows:

$$s = (2 \times 4) / (12 + 6) = 0.44 \quad (2)$$

Benedicte Pierrejean [5] evaluated 5000 text snippets generated by Nomao’s NLGs and found rather satisfactory results: 0.13 on average.

Although this indicator makes it possible to account for the diversity of Nomao’s NLGs, it gives no information on how end-users read them and how they understand them.

3. Theoretical and Methodological Framework

In the previous section, we described the functioning of Nomao’s NLGs. We have seen that this functioning is divided into three steps serve to ensure the coherence of macrostructure, microstructure and surface generated texts. We also stressed that this functioning covers a part of randomness which aims to improve the variety of content of the texts produced by Nomao’s NLGs. We have said that the quality of this variety is important for Nomao insofar as:

- it should improve Nomao in the SERP ranking;
- it should ensure that Nomao’s users do not perceive that the contents available on this site are produced by a machine.

We added that the quality of the texts produced by Nomao’s NGLs have so far been a quantitative evaluation based essentially on a measure of similarity (*i.e.*, the Dice coefficient). Finally, we pointed that, if the quality of the texts produced by Nomao’s NLGs appears satisfactory from the point of view of their objective variety (cf. the Dice coefficient). We know nothing of their quality from the point of view of the subjects who read them¹.

Also, this third section aims at defining a theoretical and methodological framework for evaluating, from user point of view, the quality of texts generated by Nomao’s NLGs.

To achieve this goal, we will first show that cognitive load theory is adequate for identifying the linguistic complexity of text from the user perspective. We then will present different methods—qualified as “on-line” and “off-line”—which are generally used to evaluate:

- cognitive and emotional load involved in reading activities;
- levels of understanding of textual documents.

¹Note that it is possible that this objective variety is an effect on how users perceive the quality of the texts generated by Nomao’s NLGs. We would also point out that it would be interesting to extend our work in seeking to better identify and understand the potential relationship.

Finally, we will present the assumptions and variables that we have used to evaluate, from the user perspective, the quality of texts generated by Nomao's NLGs.

3.1. Cognitive Load Theory

The cognitive load theory is often used to evaluate reading activities whether they are linear or non-linear [6] [7]. According to Sweller [8], there are three types of cognitive load.

Extrinsic load is related to the volume of information conveyed by the media and the manner in which this information is presented to the reader. It refers both to the content (*i.e.*, the number of information items) and to the form (*i.e.*, information formatting) of a particular text. As such, it is possible to have the extrinsic load level vary by manipulating either the quantity or the quality of the information. **Intrinsic load** is directly related to the task. The only way to reduce it is to delete some of the task's elements. This load varies depending on the user's expertise level. Finally, **essential load** depends on the ease of integration of the information in Long Term Memory (LTM). This load enables the user to acquire new knowledge.

Thus, regardless of the intrinsic characteristics of the reader, a text may create a more or less important extrinsic load depending on the way it is written. Consequently, depending on its syntactic and semantic structure, a text may increase or decrease the intrinsic and essential loads that are necessary for its understanding.

More specifically, understanding in the course of reading implies that the reader will internally build different levels of text representation through the repetition of a number of construction and integration cycles [9]. During the construction stage, the reader uses various text meaning production rules so as to organize components of the text into an association network that forms an intelligible text base. The integration phase consists in strengthening the construction process by selecting items that appear to the reader as the most relevant and by inhibiting those that seem the least appropriate.

Thus, the way a reader understands a text depends on his prior knowledge of the words that make up the text, on their graphic form, on their meaning but also on the grammatical and syntactic forms that structure the text. In other words, the understanding of a text by a reader depends on its linguistic complexity that is on the difficulties that a reader will experience while decoding the message conveyed by the text and inferring the significations that are associated with it [10].

From the reader perspective, a text with a strong linguistic complexity requires expensive processing in the Working Memory (WM). Because it involves a higher activation of the WM, increasing the linguistic complexity of a text may result in reducing the resources that the reader can allocate for storing information in LTM and thus hinder the understanding of the text [8] [11] [12] [13].

Two types of methods are generally used to evaluate both subjectively and ob-

jectively the cognitive load: on-line methods (e.g., eye movement recording) and off-line methods (e.g., survey that evaluate mental and emotional load and recall task techniques).

3.2. On-Line Methods: Eye Movement Recording

Eye movement recording is a good method to observe reading activities since it makes it possible to objectively measure the cognitive load involved during the processing of a text. Therefore, eye movement recording is a good way to evaluate the linguistic complexity of a text from reader perspective. Indeed, measuring the amount of cognitive load involved in a reading task is a way to evaluate the difficulties that the reader can meet as he reads the text. Therefore, it is an adequate way to measure the level of linguistic complexity from the reader's point of view (cf. **Table 1**).

Past works on eye movement analysis in reading situations generally show that eye movements are composed of a sequence of saccades and fixations that form three sequences which can in turn be associated with two different states [14] [15] [16]:

- at state n , the reader changes the focus of his attention from the word he was fixing (sequence 1) and moves his attention to a new word to be fixed (*i.e.*, saccade stage during which no information is processed; sequence 2);
- at state $n + 1$, the reader focuses his attention on the new word to be fixed (*i.e.*, fixation stage during which information is processed; sequence 3).

Many research works on reading activities show that the more a word is simple, the more its recognition can be achieved peripherally during the fixation stage [17] [18]. In this way, the short time span between the preparation of the saccade before state $n + 1$ (see sequence 1 to state n) and its implementation (see sequence 2 of the state n) may be sufficient to allow the reader to understand the word corresponding to $n + 1$ (see attachment sequence 3). Therefore, during this short time span, the saccade can be reprogrammed: the reader ignores the old $n + 1$ fixation state that was previously programmed during the state n and reprograms the saccade to move to a new $n + 1$ state. Thus, based on the work of O'Regan [17] and Rayner [18], it is possible to say that the more a text is made of simple words, the larger are the saccades. Plus, the saccades would also tend to be more numerous (we shall see below that theoretically the simpler a text, the

Table 1. Linguistic complexity indicators.

	Low linguistic complexity	High linguistic complexity
Number of saccades	High	Low
Duration of saccades	Long	Short
Amplitude of saccades	Great	Small
Number of fixations	High	Low
Fixation time	Short	Long

more fixations will occur; thus increasing the number of saccades) and to be longer in time (because the saccades would be larger²).

Equally, many research works in the field of air traffic control show that air traffic controllers perform a high number of saccades when they examine a complex situation (see the phenomenon of “attentional tunneling”; [19] [20] [21] [22]). Furthermore, Stein [23] found that the duration of saccades tends to decrease if the situation examined by the air traffic controller is more complex. Still, this last observation is has come under debate [24]. Consequently, a text showing strong linguistic complexity should lead the reader to realize fewer saccades and these saccades should be of short duration.

Furthermore, it has been clearly demonstrated in research about air traffic control and reading, that an increase in the complexity of an air control situation or of a text tends to decrease the amplitude of the saccades performed by the controller or the reader [23] [25] [26] [27]. In other words, in reference to O’Regan [17] and Rayner [18], in sequence of complex words a reader cannot recognize as well words in a peripheral manner. Thus, the reader cannot ignore and reprogram fixation steps; he cannot perform ample saccades as well.

We can conclude from these various research results that the linguistic complexity of a text, because it involves a greater cognitive load for the reader, can be objectively observed by checking the number of saccades, their duration and amplitude.

As far as fixations are concerned, research in air traffic control show that they are less frequent when the situation is less complex [23] [28]. Similarly, a text with a strong linguistic complexity should lead the reader to perform fewer fixations.

Concerning the duration of the fixations, it should be noted that:

- frequent and common words are subject to shorter fixation periods [29] [30] [31] [32] [33];
- the words which can be easily predicted from their context are also subject to shorter fixation periods [34] [35] [36] [37] [38];
- the words that are less predictable are subject to longer fixation periods [37] [39] [40] [41];
- abstract and/or ambiguous syntactic forms can generate longer fixation periods [42] [43].

In short, these studies results show that if the linguistic complexity of a text increase (which implies a greater cognitive load for the reader), it can be objectively observed through a lower number of fixations and an increase in their duration.

3.3. Off-Line Methods: Emotional and Cognitive Load Questionnaires, Recall Tasks

As explained above, the study of eye movement is an interesting method to objectively observe the level of cognitive load involved in a reading task. However,

²Note that the latter is widely discussed.

analysis of eye movement does not give information on how well or how much this cognitive load is subjectively perceived by the user.

This is why eye movement recording is often completed with a cognitive load evaluation questionnaire. Gerjets, Scheiter and Catrambone [44] propose a shorter version of the NASA-TLX questionnaire, which evaluates the three aspects of cognitive load [8] and the stress of the user:

- the level of mental activity required to read the document (this aspect evaluates the requirements of the task; see the intrinsic load);
- the level of mental work required to understand the information displayed in the document (this aspect evaluates the effort involved in achieving the task; see the essential load);
- the effort required to navigate inside the document (*i.e.*, to find and retrieve relevant information; this aspect evaluations navigation demands; cf. extrinsic load);
- the level of stress experienced during the reading task.

This version of the NASA-TLX questionnaire, which is designed to evaluate subjective cognitive load involved in reading hypertext documents, is well suited for evaluating the cognitive load involved in reading linear texts.

In addition, some studies highlight that the emotional load, whether positive or negative, has effects on cognition [45]. Correspondingly, Raufaste, Mariné and Eyrolle [46] show that positive emotions often allow decision-takers to perform more complex and effective cognitive processing. Conversely, negative emotions always tend to alter and degrade decision-takers' cognitive processing. Furthermore, in reference to the work of Isen [47], it seems that negative emotions tend to lead to attentional focus, which, by encouraging deep processing, consumes a lot of resources. In contrast, positive emotions tend to lead to some attentional aperture and improve synthesis ability (the ability to integrate knowledge and the information disclosed and to infer relationships among disparate ideas to build a solution).

Thus, increasing the linguistic complexity of a text may reduce the resources that the reader allocates for storing the information in LTM by increasing the workload of the WM [11] [12] [13]. However, this complexity may also have the effect of improving the emotional load involved in the reading. In this way, by inducing positive emotions in the mind of the reader, a difficult text may ultimately have the effect of enabling the reader to perform more cognitively complex and effective processing [46]. Thereby increasing the complexity of the text may, to some extent, facilitate its understanding and memorization by encouraging attentional aperture and improving the reader synthesis ability [47].

Thus, measuring the emotional load involved in reading seems to be an adequate manner to better understand the role of the linguistic complexity of a text on its understanding. More precisely, in the case of the understanding and memorization of a text describing a business (BDT—Business Describing Text), aspects such as reader confidence (which is one Plutchik's [48] eight fundamental emotions) and reader interest (related to anticipation; also part of Plutchik's

[48] fundamental emotions) should be taken into account. Confidence and interest are indeed two important dimensions in processes like information dissemination and appropriation on the market [49] [50] [51] [52].

Finally, to objectively determine whether the increase of linguistic complexity of a text rather has negative or positive effects on its understanding, recall tasks and response time measurements are generally used [53] [54]. These two types of measurement are carried out by asking the reader questions about the form and the content of the text and by monitoring the time spent in answering these questions. Although this measure is subject to debate [55], pupil diameter measurement during the response time can help to indicate the level of cognitive load involved during the processing and the elaboration of the answer: the larger the diameter, the higher the cognitive load.

3.4. Assumptions and Variables

As explained below, the experimentation consisted in asking participants to read the two texts below: Text 1 (T1) entitled “Angelina” was generated by Nomao’s NLGs (*i.e.*, a machine-generated text) and Text 2 (T2) entitled “Chez Janou” was written by humans.

Texts 1 and 2 have approximately the same number of characters (c) and words (w)—respectively 813 c and 159 w versus 767 c and 150 w . In addition, the number of characters per word (c/w) is identical in the two texts (cf. **Table 2**). From a quantitative viewpoint, texts 1 and 2 are of similar complexity.

Nonetheless, if one looks at the syntactic structure, text 1 is less complex than text 2. Text 1 has 11 sentences (s) whereas text 2 has 8 sentences. Its average number of words per sentence (w/s) is lower than text 2 (14.45 w/s for T1 against 18.75 w/s for T2). In addition, text 1 contains 5 commas (v) and 4 coordinating conjunctions “and” (cc_{et}) whereas text 2 has 10 v and 8 cc_{et} . The average number of commas per sentence (v/s) and the average number of cc_{et} per sentence (cc_{et}/s) are strongly higher in text 2 (0.45 v/s and 0.36 cc_{et}/s in T1 versus 1.25 v/s and 1 cc_{et}/s for T2).

Table 2. Quantitative evaluation of text 1 (T1) and text 2 (T2) complexity.

	T1	T2
Number of characters (c)*	813	767
Word Count (w)	159	150
Average number of characters per word (c/w)	5.11	5.11
Number of sentences (s)	11	8
Average number of words per sentence (w/s)	14.45	18.75
Number of commas (v)	5	10
Average number of commas per sentence (v/s)	0.45	1.25
Number of coordinating conjunctions “et” (cc_{et})	4	8
Average number of cc_{et} per sentence (cc_{et}/s)	0.36	1

*Not including spaces.

From a qualitative viewpoint, text 1 also appears less complex than text 2. Quantitatively, text 1 is usually composed of short sentences. Contrary to text 2, text 1 does not have any subject noun phrase. It does not have any long or detailed adverbial phrases either.

Moreover, unlike text 1, text 2 has a socio-semantic coherence that enriches its macrostructure on the semantic level. In reference to socio-cognitive approaches in human and social sciences [56], the socio-semantic coherence of a text can be defined as the organized set of socially shared representations that are conveyed by the text and that should enable the reader to construct a stereotypical mental picture of its referent. Text 2 is structured around various social representations that are culturally associated with a particular region of southern France: Provence. More precisely, the socio-semantic coherence is built from the following groups of words: “very good bistro from the Marseille planet”; “Marcel Pagnol” (name of a local author); “the Provence cuisine is tasty and generous”; “atmosphere is both thrilling and friendly”; “the owner is very kind and the service is courteous”; “charming terrace where you can chill down with friends and enjoy a Pastis” (Pastis is a local drink); “to eat without breaking the bank”; “dishes from Provence”).

Conversely, given the functioning of Nomao’s NLGs, text 1 does not include any socio-semantic coherence. It does not convey any socially shared representations thus allowing readers to build a stereotyped mental image of its referent. Indeed, it is mainly composed of pieces of factual and practical information written with simple syntactic and semantic structures.

Therefore, from both a quantitative and qualitative viewpoint, text 1 displays less complex syntactic and semantic structures than text 2. However, text 2 is stylistically more elaborated than text 1 and it offers stereotypical representations of its referent in the sense that it refers to a set of socially shared representations that are culturally associated with the south-east region of France (Provence). Also, the socio-semantic coherence of text 2 is likely to produce positive emotional charges in the reader in terms of both confidence and interest. Text 2 is thus likely to foster text’s understanding and macrostructure’s memorization.

Referring to the work we have presented above, we therefore make the following two general assumptions (gA):

- gA₁: reading a text generated by Nomao’s NLGs (*i.e.*, the text 1 “Angelina”) involves—both subjectively and objectively—a lower cognitive load than the text written by humans (*i.e.*, text 2 “Chez Janou”);
- gA₂: because it involves a higher positive emotional charge, the text written by humans is likely to foster a better understanding than the text generated by Nomao’s NLGs.

Also, we propose to operationalize gA₁ and gA₂ with the four specific assumptions (sA) presented below:

- sA₁: reading a text generated by Nomao’s NLGs objectively involves less cognitive load than reading a text written by humans;
- sA₂: reading a text generated by Nomao’s NLGs subjectively involves a lesser cognitive load than reading the text written by humans;

- sA₃: reading a text generated by Nomao's NLGs subjectively involves a lesser positive emotional charge than reading a text written by humans;
- sA₄: a text generated by Nomao's NLGs is less understandable than a text written by humans.

4. Data, Equipment and Method

In this section, we describe the data, equipment and method that we used to test the hypothesis presented above.

To do this, we first describe the main characteristics of the sample population, then the experimental equipment and data processing techniques that we used in our experiments. Finally, we present the experimentation we designed and whose goal is to evaluate the performance of Nomao's NLGs from the users point of view.

4.1. Population Sample

Our experiment was conducted with 28 students from the University of Montreal. These students were not paid. 7 of them are male (60.7%) and 11 are female (39.3%). The average age is 24 with a standard deviation of 5.8. The median age is 23.

4.2. Equipment and Data Processing Software

For eye-movement recording, we used a Tobii TX 300 screen, and we set the sample rate to 60 Hz (*i.e.*, a sample is released every 17 milliseconds). The maximum angle between two fixations was set to 0.5 degrees. For data recording, we used the Tobii 3.2 software. For data processing, we used the Excel and SPSS software.

Given that each participant was exposed to two experimental situations (*i.e.*, the text 1 "Angelina" and the text 2 "Chez Janou"), we performed within-subject. The advantage of within-subjects ANOVA compared to between-subjects ANOVA is that it avoids measuring effects that are due to the characteristics of the population samples. Its disadvantage is that the measured performance may be contaminated by the order in which the tasks are performed. To avoid this, we defined two groups of 14 students (G1 and G2) who made each experimentation in a different order.

In addition, we used the Friedman test for all ordinal data available (see the cognitive and emotional questionnaires). The Friedman test is an alternative to the within-subject ANOVA in the case of ordinal data: it is a non-parametric test developed to perform one-factor within-subject experiments.

4.3. Experimentation

Our experiment consists in presenting texts 1 and 2 to each participant for a period of 90 seconds (cf. **Table 3**).

After each text display, a series of eight closed questions on key informational units of each text are asked to each participant. Once these recall tasks have been

Table 3. Overview of machine-generated text (T1) and human-written text (T2).

Text 1 (T1): “Angelina”	Text 2 (T2): “Chez Janou”
<p>Translation: “Angelina” is a restaurant located in the 1st district of Paris. It is specialized in French cuisine. There, you will taste seafood and salads. This restaurant includes a café and proposes take-away food. Angelina accepts credit card as mean of payment. If you wish to go there by public transportation, the metro station “Tuileries” is 20m away and the metro station “Pyramides” is 600 m away. Your opinion about Angelina: twenty-three reviews refer to the service, 3 are positive versus 20. The way in which the restaurant was decorated pleased 3 guests (“refined decoration”, “aspetto elegante” and “beginning charm”). No negative comment about the breakfast was found. On a total of 7, 3 internautes liked the prices. The desserts were favorably mentioned in 3 reviews (“great hot chocolate”, “delicious chocolate”, “best hot chocolate”).</p> <p>In french: “Angelina” est un restaurant situé à Paris dans le 1er arrondissement. Il est spécialisé dans la cuisine française. Vous pourrez y déguster des fruits de mer et des salades. Ce restaurant dispose d’un café et propose des plats à emporter. Angelina accepte la carte de crédit comme moyen de paiement. Si jamais vous choisissez de vous y rendre en transports en commun, vous pourrez sortir à la station Métro Tuileries qui est située à 20 mètres ou à la station Métro Pyramides qui est située à 600 mètres. Votre opinion sur Angelina: vingt-trois avis évoquent les prestations, trois sont positifs contre vingt. La manière dont a été décoré le restaurant a séduit 3 hôtes (“décor raffiné”, “aspetto elegante” et “charme du début”). Pas de commentaire négatif relevé sur le petit-déjeuner. Sur 7, 3 internautes ont apprécié le prix. Les desserts ont été mentionnés favorablement dans trois avis (“super chocolat chaud”, “chocolat délicieux” et “meilleur chocolat chaud”).</p>	<p>Translation: Come discover Chez Janou near the place des Vosges in the 4th district of Paris. A very good bistro from the Marseille planet which has been decorated for Marcel Pagnol’s fans and where the Provence cuisine is tasty and generous. The atmosphere is both thrilling and friendly. The owner is very kind and the service is courteous. Reservation is mandatory in this very busy place. Not far away from the Marais neighbourhood, the restaurant Chez Janou has an ideal location and has a charming terrace where you can chill down with friends and enjoy a Pastis or else... At lunch time, a 14.50€ menu will allow you to eat without breaking the bank, and with à la carte service, you will discover various dishes from Provence with little more than 30 €... Cold ratatouille, anchoïade and tapenade, Saint-Nectaire toasts, Provence stuffed vegetables, grilled bass with pesto, spelt and scallops risotto...</p> <p>In french: Venez découvrir Chez Janou près de la place des Vosges dans le 4^{ème} arrondissement de Paris. Un très bon bistrot de la planète Marseille spécialement décoré pour les fans de Marcel Pagnol où la cuisine provençale y est savoureuse et généreuse. Ambiance à la fois électrisante et sympathique. Le propriétaire est bien gentil et le service est courttois. La réservation est indispensable dans ce lieu très fréquenté. À deux pas du Marais, le restaurant Chez Janou jouit d’un emplacement idéal et possède une terrasse pleine de charme, où il fait bon se poser entre amis pour un Pastis ou autre... À midi, une formule à 14.50€ permet de se restaurer sans se ruiner, et à la carte, un peu plus de 30€ pour découvrir différents plats aux saveurs provençales... Assiette de ratatouille froide, anchoïade et tapenade, Toastades de saint-nectaire, Farcis provençaux, Bar grillé au pistou, Risotto d’épeautre et ses Saint-Jacques...</p>

performed, the participants answer an emotional load evaluation questionnaire (2 items: interest and confidence levels on a 1-to-9 Likert scale; see section 3.3) and a cognitive load evaluation survey (4 items: intrinsic, essential, extrinsic loads and stress levels on a 1-to-100 Likert scale; see section 3.3). At the end of the each experiment, we asked different open questions to the participants so as to make them verbalize the perceptions they had of each text.

Recall that both 14 students groups G1 and G2 carried out the experiments in a different order to avoid contamination effects that are usually associated with within-subject experiments. Thus, G1 started with text 1 and G2 started with text 2.

5. Results

In this section, we present the results of the experimentation. Recall that, in general, these results are intended to test specific hypotheses sA1, sA2, sA3 and sA4 in order to allow the testing of general hypotheses gA1 and gA2 that we outlined in Section 3.4. So, the following subsections describe the effect of the text on eye behavior, the effect of the text on subjective cognitive and emotional loads, the effect of the text on understanding. The final subsection discusses all the results.

5.1. Text Effect on Eye Behavior

As illustrated in **Table 4**, from eye-movement analysis viewpoint, it appears that reading the text generated by Nomao's NLGs (*i.e.*, the text 1 "Angelina") involves an objective cognitive load significantly lower than reading the text written by humans (*i.e.*, text 2 "Chez Janou").

Indeed, although the difference between the numbers of saccades that participants present when reading texts 1 and 2 is not significant ($p > 0.05$), the average length of the saccades is significantly longer when reading text 1 than for text 2 (37.47 milliseconds (ms) for T1 against 35.86 ms for T2; $F(1,27) = 11.526$; $p < 0.01$).

Moreover, our results also show that the average vertical amplitude of the saccades is almost significantly higher ($p = 0.058$) during the reading of text 1 than during the reading of text 2 (47.70 pixels for T1 versus 42.73 pixels for T2).

In addition, although the difference between the numbers of fixations that participants perform during the reading of texts 1 and 2 is not significant ($p > 0.05$), the average duration of fixations is significantly shorter when reading text 1 than for text 2 (196.55 milliseconds for T1 versus 205.57 ms for T2; $F(1,27) = 11.951$, $p < 0.01$).

5.2. Text Effect on Subjective Cognitive and Emotional Loads

If reading text 1 implies a lesser objective cognitive load than reading text 2, it appears that from a subjective viewpoint, our participants do not perceive this difference (cf. **Table 5**).

Table 4. Evaluation of eye behavior for text 1 (T1) and text 2 (T2).

	T1	T2	Significance
Number of saccades	417.61 s*	405.46 s	$F(1,27) = 3.162$, $p = 0.870$
Average duration of saccades	37.47 ms*	35.86 ms	$F(1,27) = 11.526$; $p = 0.002$
Number of fixations	352.43 f*	347.04 f	$F(1,27) = 1.780$, $p = 0.193$
Average duration of fixations	196.55 ms	205.57 ms	$F(1,27) = 11.951$; $p = 0.002$
Average saccades vertical amplitude	47.70 px*	42.73 px	$F(1,27) = 3.922$, $p = 0.058$
Average saccades horizontal amplitude	22.84 px	19.86 px	$F(1,27) = 2.279$, $p = 0.143$

*Unit: saccade (s); millisecond (ms); fixation (f); pixel (px).

Table 5. Evaluation of subjective cognitive and emotional loads for text 1 (T1) and text 2 (T2).

	T1	T2	Significance
Cognitive load	1.54 r*	1.46 r	$\chi^2(1) = 0.143, p = 0.705$
Intrinsic load	1.46 r	1.54 r	$\chi^2(1) = 0.182, p = 0.670$
Essential load	1.54 r	1.46 r	$\chi^2(1) = 0.167, p = 0.683$
Extrinsic load	1.59 r	1.41 r	$\chi^2(1) = 1.087, p = 0.297$
Stress	1.52 r	1.48 r	$\chi^2(1) = 0.053, p = 0.819$
Emotional	1.32 r	1.68 r	$\chi^2(1) = 5.000, p = 0.025$
Interest	1.3 r	1.7 r	$\chi^2(1) = 6.368, p = 0.012$
Confidence	1.38 r	1.63 r	$\chi^2(1) = 2.882, p = 0.090$

*Unit: rank (r).

Our results actually show no statistical significance ($p > 0.05$) in how the participants evaluated the cognitive loads involved in reading texts 1 and 2 and in their different dimensions (*i.e.*, intrinsic, essential, extrinsic loads and stress; see Section 3.3).

In contrast, our results show that participants experienced an emotional load significantly less positive when reading text 1 than for text 2 (rank 1.32 in the case of T1 versus rank 1.68 for T2; $\chi^2(1) = 5.000; p < 0.05$).

More precisely, it appears that interest in text 1 is significantly lower than for text 2 (rank 1.3 in the case of Q1 versus rank 1.7 in the case of T2; $\chi^2(1) = 6.368; p < 0.05$). Furthermore, although this result is not significant ($p = 0.090$), participants also seem to give a confidence level lower than the text 1 text 2 (rank 1.38 in the case of T1 against rank 1.63 in the case of T2).

5.3. Text Effect on Understanding

Finally, as shown in **Table 6**, it seems that participants find it easier to understand text 2 than text 1.

Although the results for recall tasks are very similar (8.01/10 for T1 versus 7.81/10 for T2; $F(1,27) = 0.279; p > 0.05$), the results concerning average response time and average pupil diameter show statistically significant differences.

As illustrated in **Table 7**, it should be noted that the questions about the text 1 have an average of 93.62 characters per question (c/q) and 20.62 words per question (w/q) versus 105.37 c/q and 23.75 w/q for the questions about text 2.

Although the questions about text 1 have fewer characters and words than the questions about text 2, participants significantly take longer time to read and answer the questions about text 1 (6.140 ms in the case against T1 versus 4.940 ms in the case of T2; $F(1,27) = 16.742, p < 0.001$). Moreover, the average pupil diameter of participants is significantly larger during the answering of questions about text 1 than for text 2 (3.094 millimeters for T1 versus 3.057 mm for T2; $F(1,27) = 6.320; p < 0.05$), which also suggests difficulty.

Table 6. Evaluation of recall tasks performance, average response time and average pupil diameter for text 1 (T1) and text 2 (T2).

	T1	T2	Significance
Recall task	8.01/10*	7.81/10	$F(1,27) = 0.279, p = 0.602$
Average response time	6140 ms	4940 ms	$F(1,27) = 16,742; p = 0.000$
Average pupil diameter	3.094 mm*	3.057 mm	$F(1,27) = 6.230, p = 0.019$

*Unit: score of 10 (/10); millimeter (mm).

Table 7. Quantitative evaluation of the complexity of the questions about text 1 (q T1) and text 2 (q T2).

	q T1	q T2
Number of characters (c)	749	843
Word Count (w)	165	190
Number of questions (q)	8	8
Average number of characters per question (c/q)	93.62	105.37
Average number of words per question (w/q)	20.62	23.75

5.4. Discussion

Under sA_1 , we found that the text generated by Nomao's NLGs (*i.e.*, the text 1 "Angelina") objectively involves a less important cognitive load than the text written by humans (*i.e.*, Text 2 "Chez Janou").

However, our results do not confirm sA_2 since we did not find significant differences between the level of subjective cognitive load involved in the reading of text 1 (machine-generated text) and the level of subjective cognitive load involved in the reading of text 2 (written by humans).

sA_3 is validated since our results show that reading text 1 (machine-generated) subjectively implies a significantly less positive emotional charge than reading text 2 (written by humans).

Our results also confirm sA_4 : we have found that participants significantly better understand the human-written text than the machine-generated text. Note however that this difference is not visible in the recall tasks scores, only the response time and pupillary diameters back up this assumption.

On a general level, it appears that Nomao's NLGs fulfills its function since, during the questions-verbalization stage done at the end of each experiment, all participants stated that the machine-generated text was correctly written that it seemed to be written by competent humans. Thus, we can claim that the micro-planning stages are properly carried out by Nomao's NLGs since the syntactic and semantic structures of text 1 objectively posed no reading difficulties to the participants.

Experiments also show that participants do not subjectively perceive a lower cognitive load while reading the machine-generated text (which refutes sA_2). This might be due to the fact that text 1 is not structured around a set of socially shared representations as in the case of text 2 which is written by humans; these

socially shared representations generally produce positive emotions in the reader and thus facilitate the general understanding of the text.

Indeed, the machine-generated text is composed of a sequence of practical and factual information presented in the form of simple syntactic and semantic structures that are partly organized at random. Moreover, although it is subject to some kind of organization during macro-planning, the macrostructure of the machine-generated text does not have any particular socio-semantic coherence like in human-written texts. Thus it seems that this lack of socio-semantic coherence in the machine-generated text explains why the participants experienced a more positive emotional charge and a better understanding when they read the human-written text.

From a scientific perspective, our results highlight the importance of socio-semantic coherence in text-based communication produced by NLGs. This is worth noticing since socio-semantic coherence is generally not considered in the modeling of reading activities in cognitive science research. For example, Gernsbacher [57] explains that the matching processes allowing readers to build a mental representation of a text depend on the text's referential, temporal, spatial, structural and causal coherence which all differ from socio-semantic coherence. Yet, it is this latter form of coherence that is lacking in the texts generated by Nomao's NLGs.

Finally, we wish to underline that our results are exploratory and should therefore be extended over a larger amount of text and a larger population sample in order to control the level of replicability as well as the mediating effect that may be produced by individual variables that we did not consider here. It is possible that the dual task paradigm [58] could better explain the effect of the text on the cognitive and emotional loads involved in the reading task and in the understanding of the text.

6. Conclusions

In short, the case study that we proposed in this paper is interesting for two main reasons.

The first is that it emphasizes that the objective indices for measuring the quality of content generated by NGLs such as the Dice coefficient are not sufficient. Our article allows to point out the need to evaluate this quality from the point of view of users who read these contents in order to truly identify and understand the communicative performance of an automatically generated text. We then showed how on-line (*i.e.*, eye tracking) and off-line (*i.e.*, questionnaire) measurements on cognitive and emotional charge can be used to perform this type of evaluation.

The second reason is that the case study we have proposed highlights the role of socio-semantic coherence in textual communication. Our work thus points to a dimension of textual communication that is, to our knowledge, not sufficiently considered in the field of NLG. Our results show that for a text to be understood and memorized, it must not only make it easy to read, that is to say coherent

from the point of view of its macrostructure, microstructure and surface. Our study shows that for a text to succeed in transmitting easily the “communicative intention” [10] of which it carries, it is also necessary that the representations that it conveys be endowed with socially shared meanings. Also, it seems to us that in the era of big data, one of the main challenges of NLG might be to seek new artificial intelligence technology to automate the production of this socio-semantic coherence.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript.

References

- [1] Moulrier Boutang, Y. (2012) Cognitive Capitalism. Polity Press, Cambridge.
- [2] Parasie, S. and Dagiral, E. (2013) Data-Driven Journalism and the Public Good: “Computer-Assisted Reporters” and “Programming-Journalists” in Chicago. *New Media & Society*, **15**, 853-871. <https://doi.org/10.1177/1461444812463345>
- [3] Mitkov, R. (2003) The Oxford Handbook of Computational Linguistics. Oxford University Press, Oxford.
- [4] Cardon, D. (2013) In the Mind of Page Rank. A Study of Google’s Algorithm. *Réseaux*, **1**, 63-95.
- [5] Pierrejean, B. (2013) Amélioration d’un système de génération automatique de texte pour les commerces locaux. Master Thesis, University Toulouse Jean Jaurès, Toulouse.
- [6] Amadiou, F., Tricot, A. and Mariné, C. (2009) Prior Knowledge from Learning in a Non-Linear Electronic Document: Disorientation and Coherence of the Reading Sequences. *Computer in Human Behavior*, **25**, 381-388.
- [7] Amadiou, F., Tricot, A. and Mariné, C. (2010) Interaction between Prior Knowledge and Concept-Map Structure on Hypertext Comprehension, Coherence of Reading Orders and Disorientation. *Interacting with Computers*, **22**, 88-97. <https://doi.org/10.1016/j.intcom.2009.07.001>
- [8] Sweller, J. (1988) Cognitive Load during Problem Solving: Effects on Learning. *Cognitive Science*, **12**, 257-285. https://doi.org/10.1207/s15516709cog1202_4
- [9] Kintsch, W. (1988) The Role of Knowledge in Discourse Comprehension: A Constructive-Integration Model. *Psychological Review*, **95**, 163-182. <https://doi.org/10.1037/0033-295X.95.2.163>
- [10] Sperber, D. and Wilson, D. (1986) Relevance. Communication and Cognition. Blackwell, Oxford.
- [11] Gaonac’h, P. and Larigauderie, D. (2000) Mémoire et fonctionnement cognitif: La mémoire de travail. Armand Colin, Paris.
- [12] Just, M.A. and Carpenter, P.A. (1992) A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review*, **99**, 122-149. <https://doi.org/10.1037/0033-295X.99.1.122>
- [13] Oakhill, J.V. Cain, K. and Yuill, N. (1998) Individual Differences in Children’s Understanding Skill: Toward an Integrated Model. In: Hulme, C. and Joshi, R.M., Eds., *Reading and Spelling: Development and Disorders*, Erlbaum, Mahwah, 343- 367.
- [14] Fisher, B. (1986) The Role of Care in Visually Guided Eye Movements in Monkey

- and Man. *Psychological Research*, **48**, 251-257. <https://doi.org/10.1007/BF00309089>
- [15] Fisher, B. and Breitmeyer, B. (1987) Mechanisms of Visual Attention Revealed by Saccadic Eye Movements. *Neuropsychologia*, **25**, 73-83.
- [16] Morrison, R.E. (1984) Manipulation of Stimulus Onset Delay in Reading: Evidence for Parallel Programming of Saccades.. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 667-682. <https://doi.org/10.1037/0096-1523.10.5.667>
- [17] O'Regan, J.K. (1975) Constraints on Eye Movements in Reading. Unpublished Doctoral Thesis, University of Cambridge, Cambridge.
- [18] Rayner, K. (1979) Eye Guidance in Reading: Fixation Locations within Words. *Perception*, **8**, 21-30. <https://doi.org/10.1068/p080021>
- [19] Cabon, P., Farbos, B. and Mollard, R. (2000) Gaze Analysis and Psychophysiological Parameters: A Tool for the Design and the Evaluation of Man-Machine Interfaces. Technical Report No. 2000-015, EUROCONTROL Experimental Center, France.
- [20] Endsley, M.R. and Rodgers, M.D. (1998) Distribution of Attention Situation Awareness, Workload and in a Passive Air Traffic Control Task: Implications for Operational Errors and Automation. *Air Traffic Control Quarterly*, **6**, 21-44.
- [21] Tole, J.R., Stephens, A.T., Harris, R.L. and Ephrath, A.R. (1982) Visual Scanning Behavior and Mental Workload in Aircraft Pilots. *Aviation, Space, and Environmental Medicine*, **53**, 54-61.
- [22] Willems, B., Allen, R.C. and Stein, E.S. (1999) Air Traffic Control Specialist Visual Scanning II: Task Load, Visual Noise, and Intrusions into Controlled Airspace. Technical Report No. DOT/FAA/CT-TN99/23, William Hughes FAA Technical Center, Atlantic City.
- [23] Stein, E.S. (1992) Air Traffic Control Visual Scanning. Technical Report. No. DOT/FAA/CT-TN92/16, William Hughes FAA Technical Center, Atlantic City.
- [24] Willems, B. and Truitt, T.R. (1999) Implications of Reduced Involvement in Road Traffic Air Control. Technical Report No. DOT/FAA/CT-TN99/22, William Hughes FAA Technical Center, Atlantic City.
- [25] Ahlstrom, U. and Friedman-Berg, F.J. (2006) Using Eye Movement Activity as a Correlate of Cognitive Workload. *International Journal of Industrial Ergonomics*, **36**, 623-636.
- [26] Pavlidis, G.T. (1983) The "Dyslexia Syndrome" Objective and Its Diagnosis by Erratic Eye Movements. In: Rayner, K., Ed., *Eye Movements in Reading: Perceptual and Language Processes*, Academic Press, New York.
- [27] Rayner, K. (1978) Eye Movements in Reading and Information Processing. *Psychological Bulletin*, **85**, 618-660. <https://doi.org/10.1037/0033-2909.85.3.618>
- [28] Togami, H. (1984) Affects on Visual Search Performance of Individual Differences in Fixation Time and Number of Fixations. *Ergonomics*, **27**, 789-799. <https://doi.org/10.1080/00140138408963552>
- [29] Chaffin, R. Morris, R.K. and Seely, R.E. (2001) Learning New Word Meanings from Context: A Study of Eye Movements. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **27**, 225-235. <https://doi.org/10.1037/0278-7393.27.1.225>
- [30] Inhoff, A.W. and Rayner, K. (1986) Parafoveal Word Processing during Eye Fixations in Reading: Effects of Word Frequency. *Perception & Psychophysics*, **40**, 431-439. <https://doi.org/10.3758/BF03208203>
- [31] Juhasz, B.J. and Rayner, K. (2003) Investigating the Effects of a Set of Intercorrelated Variables on Eye Fixation Durations in Reading. *Journal of Experimental*

Psychology: Learning, Memory and Cognition, **29**, 1312-1318.

<https://doi.org/10.1037/0278-7393.29.6.1312>

- [32] Rayner, K. and Duffy, S. (1986) Lexical Complexity and Fixation Times in Reading: Effects of Word Frequency, Verb Complexity, and lexical Ambiguity. *Memory & Cognition*, **14**, 191-201. <https://doi.org/10.3758/BF03197692>
- [33] Williams, R.S. and Morris, R.K. (2004) Eye Movements, Word Familiarity, and Vocabulary Acquisition. *European Journal of Cognitive Psychology*, **16**, 312-339. <https://doi.org/10.1080/09541440340000196>
- [34] Ehrlich, S.F. and Rayner, K. (1981) Contextual Effects on Word Perception and Eye Movements during Reading. *Journal of Verbal Learning and Verbal Behavior*, **20**, 641-655.
- [35] Balota, D.A., Pollatsek, A. and Rayner, K. (1985) The Interaction of Contextual Constraints and Parafoveal Visual Information in Reading. *Cognitive Psychology*, **17**, 364-388.
- [36] Rayner, K. and Well, A.D. (1996) Effects of Contextual Constraint on Eye Movements in Reading: A Further Examination. *Psychonomic Bulletin & Review*, **3**, 504-509. <https://doi.org/10.3758/BF03214555>
- [37] Rayner, K., Warren, T., Juhasz, B.J. and Liversedge, S.P. (2004) The Effect of Plausibility on Eye Movements in Reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **30**, 1290-1301. <https://doi.org/10.1037/0278-7393.30.6.1290>
- [38] Ashby, J. Rayner, K. and Clifton, C.J. (2005) Eye Movements of Highly Skilled and Average Readers: Differential Effects of Frequency and Predictability. *The Quarterly Journal of Experimental Psychology Section A*, **58**, 1065-1086. <https://doi.org/10.1080/02724980443000476>
- [39] Murray, W.S. and Rowan, M. (1998) Early, Mandatory, Pragmatic Processing. *Journal of Psycholinguistic Research*, **27**, 1-22. <https://doi.org/10.1023/A:1023233406227>
- [40] Ni, W.J. (1996) Sidestepping Garden Paths: Assessing the Contributions of Syntax, Semantics and Plausibility in Resolving Ambiguities. *Language and Cognitive Processes*, **11**, 283-334. <https://doi.org/10.1080/016909696387196>
- [41] Ni, W., Fodor, J.D., Crain, S. and Shankweiler, D. (1998) Anomaly Detection: Eye Movement Patterns. *Journal of Psycholinguistic Research*, **27**, 515-539. <https://doi.org/10.1023/A:1024996828734>
- [42] Rayner, K., Carlson, M. and Frazier, L. (1983) The Interaction of Syntax and Semantics during Sentence Processing: Eye Movements in the Analysis of Semantically Biased Sentences. *Journal of Verbal Learning and Verbal Behavior*, **22**, 358-374.
- [43] Clifton, C. (2003) We Eye-Tracking Using Data to Evaluate Theories of On-Line Sentence Processing: The Case of Reduced Relative Clauses. *12th European Conference on Eye Movements*, Dundee, Scotland.
- [44] Gerjets, P., Scheiter, K. and Catrambone, R. (2004) Designing Instructional Examples to Reduce Intrinsic Cognitive Load: Molar versus Modular Presentation of Solution Procedures. *Instructional Science*, **32**, 33-58. <https://doi.org/10.1023/B:TRUC.0000021809.10236.71>
- [45] Damasio, A.R. (1995) *L'Erreur de Descartes: La raison des émotions*. Odile Jacob, Paris, 293.
- [46] Raufaste, E., Eyrolle, H. and Mariné, C. (1998) Pertinence Generation in Radiological Diagnosis: Spreading Activation and the Nature of Expertise. *Cognitive Science*, **22**, 517-546. https://doi.org/10.1207/s15516709cog2204_4

- [47] Isen, A.M. (1993) Positive and Affect Decision Making. In: Lewis, J. and Haviland, M., Eds., *Handbook of Emotions*, Guilford Press, New York.
- [48] Plutchik, R. (1980) *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, New York.
- [49] Hirschman, A. (1977) *The Passions and the Interests: Political Arguments for Capitalism before Its Triumph*. Princeton University Press, Princeton.
- [50] Yalch, R.F. and Yoshida, S. (1983) Consumer Use of Information and Confidence in Making Judgements about a New Food Store: An Attribution Theory Analysis. *Advances in Consumer Research*, **10**, 40-44.
- [51] Pieniak, Z., Verbeke, W., Vermeir, I., Bruns, K. and Olsen, S.O. (2007) Consumer Interest in Fish Information and Labelling. *Journal of International Food & Agribusiness Marketing*, **19**, 117-141. https://doi.org/10.1300/J047v19n02_07
- [52] Swedberg, R. (2010) The Structure of Confidence and the Collapse of Lehman Brothers. *Research in the Sociology of Organizations*, **30A**, 71-114.
- [53] Dee-Lucas, A. and Larkin, J.H. (1995) Learning from Electronic Texts: Effects of Interactive Overviews for Information Access. *Cognition & Instruction*, **13**, 431-468.
- [54] Kellman, P.J. and Massey, C.M. (2013) Perceptual Learning, Cognition and Expertise. *Psychology of Learning and Motivation*, **58**, 117-165.
- [55] Chen, S. and Epps, J. (2013) Automatic Classification of Eye Activity for Cognitive Load Measurement with Emotion Interference. *Computer Methods and Programs in Biomedicine*, **110**, 111-124. <https://doi.org/10.1016/j.cmpb.2012.10.021>
- [56] Wassilew, A.Z. (2014) A Social-Cognitive Approach to Semantic Analysis and Categorization Processes. *Online Journal of Applied Knowledge Management*, **2**, 172-181.
- [57] Gernsbacher, M.A. (1996) Coherence during Mapping Cues Understanding. In: Costermans, J. and Fayol, M., Eds., *Processing Interclausal Relationships in the Production and Comprehension of Text*, Erlbaum, Hillsdale.
- [58] Kellogg, R.T. (1987) Effects of Topic Knowledge on the Allocation of Processing Time and Cognitive Effort to Writing Processes. *Memory & Cognition*, **15**, 256-266. <https://doi.org/10.3758/BF03197724>



Scientific Research Publishing

Submit or recommend next manuscript to SCIRP and we will provide best service for you:

Accepting pre-submission inquiries through Email, Facebook, LinkedIn, Twitter, etc.

A wide selection of journals (inclusive of 9 subjects, more than 200 journals)

Providing 24-hour high-quality service

User-friendly online submission system

Fair and swift peer-review system

Efficient typesetting and proofreading procedure

Display of the result of downloads and visits, as well as the number of cited articles

Maximum dissemination of your research work

Submit your manuscript at: <http://papersubmission.scirp.org/>

Or contact jcc@scirp.org

