DIGITAL LITERACY IN FOREIGN LANGUAGE THROUGH TEXT MINING AND FANFICTION WRITING

Letramento digital em língua estrangeira por meio da mineração de texto e da escrita de fanfiction

Patrícia da Silva Campelo Costa BARCELLOS
Universidade Federal do Rio Grande do Sul
patricia.campelo@ufrgs.br
https://orcid.org/0000-0002-5142-4730

Eliseo Berni REATEGUI
Universidade Federal do Rio Grande do Sul
eliseoreategui@gmail.com
https://orcid.org/0000-0002-5025-9710

Eunice POLONIA
Universidade Federal do Rio Grande do Sul
profepoloniaufrgs@gmail.com
https://orcid.org/0000-0002-6852-1473

Rebecca BLACK
Universidade da Califórnia
rwblack@uci.edu
https://orcid.org/0000-0003-2412-6217

ABSTRACT: This study investigates how digital literacy in a foreign language (FL) may be supported by the use of a digital resource that can aid the processes of reading and writing. Thus, this research is based on studies by Feldman and Sanger (2006) about text mining, and on research by Black (2007; 2009) about the incorporation of a text genre typical from the Internet, the fanfiction (text based on existing media), in language learning. Through the use of the text mining resource *Sobek*, which promotes the extraction of frequent terms present in a text, the participants of this study created digital media narratives in English as a foreign language (FL). The undergraduate Brazilian students who participated in the research used the tool *Sobek* to mediate the production of fanfictions. In the proposed task, each student read a fanfiction and used the mining tool to develop graphs with recurrent terms found in the story. From the data analysis, it was observed that the use of a digital tool supported text production in the FL, and its following practice of digital literacy, as

the authors relied on the mining resource to create new fanfictions.
KEYWORDS: Digital literacy; Fanfiction; Text mining.

RESUMO: Este estudo investiga como o letramento digital em uma língua estrangeira (LE) pode ser apoiado pelo uso de um recurso digital que auxilie os processos de leitura e escrita. Assim, esta pesquisa se baseia em estudos de Feldman e Sanger (2006) sobre mineração de texto e em pesquisas de Black (2007; 2009) sobre a incorporação de um gênero de texto característico da *Internet*, a *fanfiction* (textos baseados em diversas mídias), no aprendizado de língua inglesa. Por meio do uso do recurso de mineração de texto *Sobek*, que promove a extração de termos frequentes presentes em um texto, os participantes deste estudo criaram narrativas de mídia digital em inglês como LE. Os estudantes de graduação que participaram da pesquisa utilizaram a ferramenta *Sobek* para mediar a produção de *fanfictions*. Na tarefa proposta, cada aluno leu uma *fanfiction* e usou a ferramenta de mineração para desenvolver gráficos com termos recorrentes encontrados na história. A partir da análise dos dados, observou-se que o uso de uma ferramenta digital apoiava a produção de texto na LE, e uma prática subsequente de letramento digital, pois os autores contavam com o recurso de mineração para criar novos exemplos de *fanfiction*.
PALAVRAS-CHAVE: *Fanfiction*; Letramento digital; Mineração de texto.

INTRODUCTION

Digital tools that may be effective in the process of literacy development in a foreign language (FL) can act as mediating resources for pedagogical tasks. Several scholars have established facilitative relationships among technology, language and literacy (KERN; WARSCHAUER, 2000; LANKSHEAR; KNOBEL, 2007; WARSCHAUER, 2001). Thus, it is important that teachers working in these converging areas make use of electronic resources in a meaningful way, in order to create tasks that may provide opportunities for literacy and language development in an FL. In this study, literacy is considered a process through which language learning can occur based on the development of socially meaningful reading and writing practices (BARTON; HAMILTON, 2000). Taking as a starting point the reading and analysis of the text genre fanfiction (story published on the internet, produced by fans of media symbols, represented by movies and TV, for instance), the process of literacy in an FL is examined, bearing in mind the relationship between technology and language learning. Thus, we present in this study a discussion

about cognitive and social processes related to literacy in an FL, supported by digital means.

This paper also presents a new approach to the use of text mining as a support for literacy in an FL. We investigate how a tool for text mining can assist with the processes of reading and writing fanfiction in an FL, so as to foster literacy practices. This tool operates by making automatic identification of relevant terms and their relationships in a text, presenting an image represented by a graph of such relationships. Hence, we can observe how the digital environment can be a tool to support the learning process through a task in which the text mining tool mediates written production. To this end, we investigated how the participants dealt with the process of creating fanfiction narratives using the mining tool. Then, it was observed how the digital mining device might sustain written production of a typical text genre from the web.

Feldman and Sanger (2006) define such mining processes as a method of extracting relevant information from databases. In a broader sense, data mining processes have been used in research focused on educational systems, as indicated by Romero and Ventura (2007), who state that the field of educational data mining has received great emphasis in recent years. For example, some studies have turned their attention to online educational materials (GROBELNIK; MLADENIC; JERMOL, 2002) and online educational environments (SEVINDIK; DEMIRKESER, 2010). However, in the field of FL learning, there is little research that connects literacy and language development to the process of text mining. Most studies that link language learning to text mining focus on identifying errors in written productions (FARKAS; POLTROCK, 1995; OGATA; YANA; WAKITA, 1998; SCHWIND, 1995). Distinctively, we look at the process of literacy development through the use of a mining tool as a mediator of written production to envision new methods for using technology, as well as popular culture to foster language and literacy learning in meaningful contexts.

LITERATURE REVIEW

Text mining and learning

According to Bamford (2003), visual images are becoming the dominant form of communication in teaching and learning resources presented in a variety of media and formats. Likewise, new imagery connected to technologies may present the following possibilities: the merging of multiple languages in a single support, the expansion of

forms of interactivity and the feasibility of collective productions (NOVA, 2003). Willis and Willis (2007) state that the presentation of information in a visual web can help learners process and organize information in a more structured way. Then students have the opportunity to organize information more consistently, before creating texts, for instance. These authors suggest that tasks which include visual elements (such as charts, graphs and maps) can assist students cognitively by providing a means of organizing their thoughts visually. In addition, information may be presented or organized by students more systematically with the help of visual aids.

Lemke (1997) highlights how teachers often do not prompt students to integrate visual elements with writing practices, since our teaching routine tends to be logocentric. However, Johnson (2001) reminds us modern graphical interface tends to appeal to spatial memory. Therefore, it is essential to incorporate pedagogical resources that integrate visual and spatial references in the learning process, such as graphic organizers, in order to provide students with useful pedagogical tools. As reported by Kim et al. (2004), graphic organizers include cognitive maps, semantic maps, graphs, and diagrams, among other structures. As stated by these authors, graphic organizers are visual and spatial samples designated to facilitate teaching and learning through the use of representative elements of graphs, such as nodes and key concepts. Hall and Strangman (2002) elaborate this definition by stating that these graphic organizers act as graphic presentations, reflecting relationships between facts, terms, and/or ideas. Additionally, graphic organizers combine the use of linguistic and non-linguistic modalities (MARZANO; PICKERING; POLLOCK, 2001), an aspect that grants a multimodal character to these graphical representations. Instead of linear texts, these organizers display relations among terms represented by spatial arrangements (MANOLI; PAPADOPOULOU, 2012). Furthermore, Vekiri (2002) asserts these organizers would work as scaffolding for the development of associations between concepts, helping learners produce more connections between ideas.

According to Kim et al. (2004), graphic organizers can be helpful since they portray key terms, concepts and relationships by drawing a visual idea of a text for an observer/learner. Thus, there is the opportunity of such organizers assisting students when collecting useful information that may help the processes of reading and writing. This kind of organization established by a graph, for example, could foster interpretive associations by the individual, as the learner can draw from ideas presented by the image to better understand a text or to begin a new production. In this sense, Manoli and Papadopoulou (2012) point out that organizers also have the function of working as visual maps, presenting a textual structure in the form of an image that can assist the

learner when creating a new text later. In this regard, the spatial arrangement of a text that has been turned into a graph may provide better possibilities for reading and writing (SUZUKI; SATO; AWAZU, 2008).

When creating a graphic organizer, one can depend on a previous linear written text to build a visual representation of this material. According to Kim et al. (2004), the use of graphic organizers is more effective when such an organizer is created by the learner, not by the teacher. In the present study, the software *Sobek* offers an automatic first version of a graph, which can then be edited by the student, who modifies relationships between concepts according to their reading. This mining tool identifies frequent terms in a text and presents relationships among them in the form of a graph. Such graphical representation can help students focus on specific elements of the text, in order to create narratives in an FL.

Sobek operates as follows: (1) a text is inserted into the main page, as shown in Figure 1; (2) based on this text, the tool automatically creates a base of terms through the process of text mining. These terms correspond to a set of words and terms that occur more frequently in the text (aside from articles, prepositions and other words which are not deemed relevant). However, the user has the possibility of editing the terms which are offered by the tool, in order to create new relations and undo automatic connections presented by *Sobek*; (3) when the concepts are extracted, a graph is generated with the most common terms and their relationships, as shown in Figure 2.

Figure 1: Main Page of *Sobek*
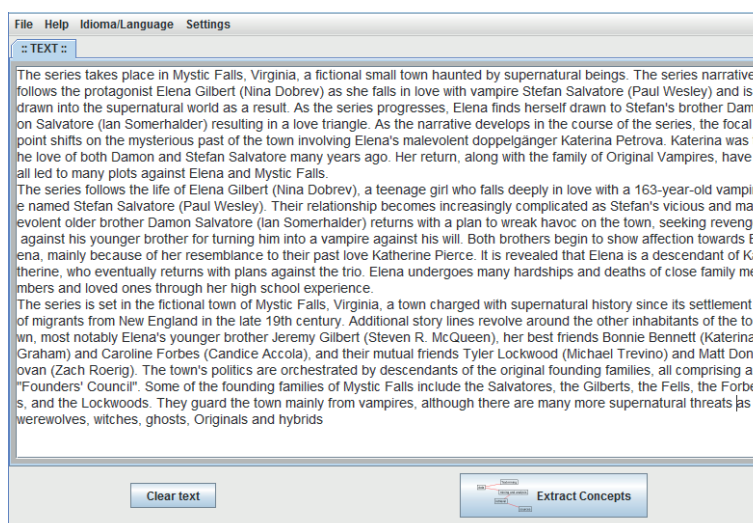
Source: Elaborated by the authors.

Figure 2: Example of a graph based on a Fanfiction about the TV show *The Vampire Diaries*
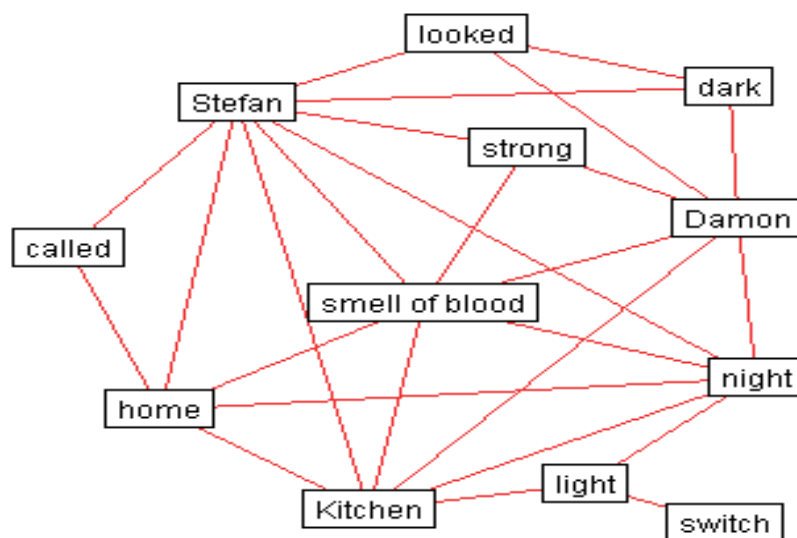
Source: Elaborated by the authors.

Such graphs may be used to help students reflect on their readings as they analyze the key terms included in the graph after having read a text. In addition, they may create and undo associations among terms and employ the tool and its graphs as support for a pre-writing stage. Specifically, in this study, students read a fanfiction of their choosing and then used *Sobek* to create a representative graph of that text. Thus, students were able to reflect on their readings as they analyzed the terms included in the graph. Moreover, they received a graph from their peers (of a different story) in order to begin composing a new fanfiction text of their own, having this image as a foundation for a new production based on media that they are interested in.

Regarding memory and mental schemes, Sternberg (2006) asserts that concepts can be organized into schemas, which are mental structures responsible for representing knowledge involving interrelated concepts in a meaningful organization. We can think of these structures as graphs representing concepts and their relationships. Moreover, still according to Sternberg (2006), knowledge can be represented as a semantic network formed by elements that represent diverse meanings and concepts. Therefore, the network provides a way for the concepts to be organized, since an individual may be able to connect nodes to a significant extent. In turn, in the present research, students can have models of graphs derived from the extraction of concepts, in order to assist them structuring a new

text, while creating the first draft.

The graph itself can act as a sketch model, which supports the beginning of a text creation. As Torrance and Galbraith (2006) argue, written production based on draft may foster a more fluid writing process, since the individual has an initial outline in which to stand. Therefore, the author would have more opportunities to organize their ideas before writing, in order to focus their attention on the act of transforming these ideas into words, the phase of creating the text itself. Furthermore, according to Graham (2006), at least three cognitive processes should be incorporated into the classroom, planning, development and edition, so as to promote effective and independent use of these procedures by the student. Therefore, in this research, the planning stage involves the generation, analysis and editing of graphs. The stage of development is related to the possibility of transforming the terms represented in the graphs in ideas, in order to elaborate an initial draft. As for the editing stage, it concerns the moment of writing the text itself, prior to the moment of sharing the production. These classifications may vary according to the study, and, in some cases, the editing stage can be considered a later stage of writing production. These phases represent pedagogical instruction centered on the writing process (ATWELL, 1998; CALKINS, 1994; TOMKINS, 1990), whereby the steps of planning or pre-writing and draft preparation are essential for students' production of texts. Therefore, the technological process based on text mining may be a useful tool in process-based approaches to FL writing instruction, in that it provides the learner with a point of departure (in the form of graphs) for crucial components of the writing process.

In the field of FL learning, there is still little research that connects communicative language practices (i.e., writing and speaking) to the process of text mining. Most studies which link language learning and text mining focus on the identification of error on written production (FARKAS; POLTROCK, 1995; OGATA; YANO; WAKITA, 1998; SCHWIND, 1995). In contrast, our goal is to understand how the text mining tool *Sobek* may be used to foster language and literacy development at a broader scale that goes beyond the identification of errors.

Literacies and Technology

In addition to employing a method of text mining, this study also incorporates interest-driven literacy practices. More specifically, undergraduate students of English as an FL read, analyzed, and wrote fanfictions, or popular culture-based texts of their choosing. It has already been investigated the role that genres of writing based on popular

culture and learners' other interests can play in language learners' development of literacy skills. For example, Black's studies (2006, 2007, 2009) demonstrated the effectiveness of writing fanfiction in the process of language learning, by showing that the informants who participated in her study developed confidence and motivation to continue writing due both to the interest and scaffolding provided by their online audience, and by writing texts based on media that was of significant interest to them and their peers.

Concerning this genre, Black (2009) defines a fanfiction as a written piece in which the stories are based on existing characters, settings, and sometimes storylines of popular media such as movies, television, and video games, to name just a few. According to this author, to create a fanfiction fans extend the plots and create new narrative possibilities between the characters of various media. Although fanfiction already existed in printed form as fanzines (JENKINS, 1992), new technologies offer these fans the opportunity to have online meetings during which authors and readers can write, exchange texts, criticize and discuss changes in their fanfictions (BLACK, 2006). Then the authors/fans have the chance to receive feedback from their readers, write collaboratively and interact with social groups who share similar interests. Besides, the procedures of production and diffusion of writing are also altered in digital means, since individuals have the opportunity to publish their own material and receive feedback from their audience. In blogging, for example, the reader's feedback is essential for the posting of new content on the webpage (LEE, 2011). In turn, the development of fanfictions is carried out in chapters, so that the continuity of the texts is linked to the response offered by readers (BLACK, 2007). Thus, profound changes occur in the relations between author and reader, which often merge the roles of producer and audience on the Internet.

Most studies have focused on learners' elective uses of technologies and the fanfiction genre in out-of-school contexts. Hence, little is known about the impact that such technologies and interest-driven forms of writing might have on students' literacy development in formal education settings. In this study, we attempt to contribute to our knowledge in this area by applying tasks that incorporate fanfiction reading and writing mediated by text mining in an academic context. According to Black (2008), many students today are accustomed to active participation in digital contexts in which communication is not only based on print language but also grounded in a varied and complex semiotic system available via digital media. Black (2008) shows how learners/fans integrate fanfiction into their daily social practices even in out-of-school contexts. The present work, however, investigates how pedagogical tasks, which include recurrent practices of the digital space, can be conducted in school context. Besides the chance of engaging

in a series of self-selected literacy practices, when there is an insertion of pedagogical activities based on real text genres, students may have the first contact with meaningful and authentic practices. Moreover, they may have more opportunities of engaging later in voluntary writing circumstances, being able to deal with different literacies.

Literacy is considered a meaningful process of social practice through which individuals can act in interactional contexts of language production (LANKSHEAR; KNOBEL, 2007). According to Scribner and Cole (1981), social practice refers to socially developed and patterned ways of using technology and knowledge to perform tasks. Hence, literacy involves the application of reading and writing skills in specific contexts of use for specific purposes. In this work, we also consider the idea of digital literacy, since individuals in the online environment also make social use of their construction of knowledge through digital means. According to Warschauer (2003), digital literacy is a process in which the user, besides having Internet connection, holds skills and comprehension of how to use the computer and the internet in a socially valuable way. Thus, we see the importance of a learning process based on literacy in language that also encompasses a concern about the possibilities of digital literacy. In this study, the practice of digital literacy in an FL is encouraged through the reading and writing of the text genre fanfiction mediated by a digital tool. According to Black (2008), the fanfiction phenomenon shows that new technologies have fostered the development of social practices and literacy practices that surpass linguistic and cultural barriers. Besides, as reported by this author (BLACK, 2008), the success of fanfiction illustrates how technological means have eased the diffusion of cultural and intellectual products developed by fans that are distributed in online communities. Hence, it is relevant that students have the opportunity to produce this genre in real contexts of use in order to increase their literacy practices in a highly digitized world.

METODOLOGY

Data Collection

The focal participants in this study are two undergraduate students of 19 and 21 years old who were in the pre-intermediate level of learning English as an FL in the Modern Languages program in a federal university of Brazil. In the classroom context, these students worked as a pair. Data were collected during five class sessions of forty-five minutes each that incorporated reading and writing fanfiction using the *Sobek* text

mining tool. Briefly, the steps involved in this learning activity were as follows:

(a) Input: students were introduced to the website Fanfiction.net, the largest existing online fanfiction archive. Next, students chose a fanfiction from the Internet to read (from Fanfiction.net or a site of their own choosing), according to their interest, and then were asked to analyze these texts. In this input stage, students had the opportunity to familiarize themselves with the central features of the online fanfiction genre and select particular texts to analyze;

(b) Analysis of texts: students used the tool *Sobek* to examine the structure of their chosen fanfictions by using the text mining software to create graphs with key terms of the text. Next, they edited the *Sobek*-produced terms (adding or deleting words according to what the students deemed important for the narrative) into their own customized graphic organizer that would be forwarded to one of their peers (this graph would then serve as the departure point for their peer's composing of a fanfiction text);

(c) Process of text production: based on a graph produced by a peer, the students created new fanfictions using the graphical aid (students exchanged their customized graphic organizers and then used their peers' graphs as a point of departure for writing their own fanfiction texts).

After they completed these tasks, there was an exploratory phase in which students' views on the learning experience were assessed. The participants responded to a questionnaire (in their mother tongue, Portuguese) in relation to their experience working with fanfiction, involving questions related to their approach to the tool *Sobek* and their experience writing fanfiction. This way, we were careful to provide an emic perspective[1] of the research, rather than just an external view (JOHNSON, 1992). Based on this emic perspective, we take into account the participant's point of view, who is not only subject to the interpretation of an external observer (the researcher). Data also consisted of graphic organizers, and screen captures of the production of their graphs and fanfiction texts (i.e., students' movements on the computer screen were recorded on video).

Questionnaires and screen captures were analyzed using qualitative coding (SALDAÑA, 2012). According to this author, a code, in qualitative research, is generally a word or short expression that symbolically indicates an attribute that stands out to a portion of linguistic or visual data. As stated by Saldaña (2012), it would be a construct generated by the researcher that symbolizes and attributes a sense of interpretation for excerpts of data. Then, the following codes are included in this work, according to the

---

[1] Emic perspective refers to the perspective of the participant (the studied social group).

codification by Saldaña (2012): a) descriptive code, which summarizes the central topic of an excerpt; b) In Vivo code, which is represented by an expression or word used by the research informant himself.

Students' texts and questionnaires were coded with an eye to recurrent as well as divergent patterns in text production and perceptions of the process. While coding the screen captures, the authors used magnitude coding to focus on: (a) the regularity with which the apprentices returned to the original text in order to edit the terms before generating a new graph; (b) the regularity with which the learners examined the graph produced by their peer while writing a new text; (c) the recurrence of terms from the graph in the new fanfiction written by the learner; (d) the presence of topics related to terms from the graph in the new fanfiction.

Analysis

We focus here on the work carried out by the students Fatima[2] and Laisa. Following the steps mentioned in the previous section, the students worked as a pair, exchanging graphs before writing their fanfiction texts. Figure 3 shows the graph produced by Fatima after reading a fanfiction about the TV show *Friends*[3]. Later, her peer Laisa wrote a new text based on this graph.

Figure 3: Graph edited by Fatima, related to a fanfiction about *Friends*

Source: Elaborated by the authors.

Table 1 presents the description of some actions taken by Fatima when editing the concepts produced by *Sobek* as she created the graph to be forwarded to her peer Laisa. The actions are presented sequentially, with the time that the user spent performing each action in the left column, and a description of the action in the right column.

Table 1: Fatima's actions when creating the graph

Time	Student's actions
15:27[4]– 15:35	Fatima chooses the fanfiction *Jumping to Conclusions* to read.
15:36 – 18:04	Fatima pastes the story and pastes onto the initial page of *Sobek*.
18:05 – 18:15	Fatima analyses the relations of terms that were established by the mining tool.
18:16 – 20:00	Fatima creates new relations among terms and removes some of the terms.
20:01 – 20:38	Fatima clicks on the icon *Create Graph* and observes the image.
20:39 – 21:12	Fatima returns to the *Concept Base* of *Sobek* and edits more concepts.

---

[4] Time concerning the beginning of the recording process. In the example, '15:27' refers to fifteen minutes and twenty-seven seconds since the beginning of data recording.

| 21:13 – 21:39 | Fatima clicks on *Create Graph* and observes the image. |
| 21:40 – 23:57 | Fatima returns to the original text on the webpage and rereads the text attentively. |

Source: Elaborated by the authors.

From these data, it is clear that the graph created by the text mining software was a salient aspect of Fatima's reading process. The graph served as a starting point for the student to begin thinking about the relationships between significant characters, concepts, and plot elements of the story. However, rather than simply relying on the graph and semantic relationships generated by the text-mining software, Fatima took an active role in restructuring the components of the graph by indicating specific relationships between characters/concepts and adding new concepts. Thus, the technology allowed the student to easily alter the graph to make it more closely aligned with her initial reading of the text. After this initial refiguring of the graph, Fatima returned to the text and reread it, presumably in order to enhance her understanding of the relations between words highlighted by the mining software. It is also worth noting that this time, Fatima appeared to make a more careful reading of the text, following each line with the scroll bar and taking more time to complete the reading than she did the first time. The questionnaire answered by Fatima also suggested that using the text mining software prompted her to engage more deeply with the text. For example, she stated "The experience was funny, because the graph sometimes had words that were linked to each other in odd ways, so it was necessary to remove them [the links]. Then, I had to go back to the original text and read it again to make new connections". Therefore, even though the *Sobek* graph provides an image to support English learners' comprehension of a text, the readers then can make their own changes to the graph to make it more aligned with their own interpretation of the text. This process can lead to moments of reflection in which the learner considers how semantic relationships are represented, not only in the *Sobek* graph but also in the text itself.

We hypothesize that this careful reflection on text elements during reading may also assist students in their writing process. Giving support to our hypothesis, Thomas and Healy (2012) emphasize that the process of rereading can foster a deeper understanding of texts. Therefore, the frequency with which the learner returned to the original text is an aspect taken into account in this work. Thomas and Healy (2012) mention that memory allows the reader to transfer what was understood during the first "encounter"

with a certain text to the following "encounters", prompting a faster reading. Indeed, we observed that subsequent readings were briefer. Also according to these authors, most beginners in a second language need to read more slowly or reread a text, for the development of a coherent mental representation of what was read.

As seen in Table 2, for this participant, there is a process of close reading after the tool is used and the graph is built, as the learner returned to the original text to support her interpretation and editing of the *Sobek*-produced graph. After this initial encounter using *Sobek*, Fatima returns to the same fanfiction canon[5] of *Friends*, selects a new story, and then goes through an identical process as she prepares the graph that will be passed on to her classmate in the next part of the activity. Basically, she engages in a process that involves going back and forth between the source text and the text-mining graph, rereading and making changes until the conceptual map represents her understanding of the story.

Table 2: Fatima's actions when creating the graph (Part II)

| Time	Student's actions
23:58 – 24:04	Fatima checks the section about *Friends* once more on the website and looks for a new fanfiction on this theme.
24:05 – 24:10	Fatima chooses the fanfiction *TOW The Central Park receipt*.
24:11 – 24:31	Fatima pastes the story and pastes onto the initial page of *Sobek*
24:32 – 24:39	Fatima analyses the relations of terms that were established by the mining tool.
24:40 – 32:56	Fatima returns to the original text on the webpage and rereads the text attentively.
32:57 – 33:00	Fatima checks the graph again.
33:01 – 35:33	Fatima returns to the *Concept Base* of *Sobek* and edits more concepts.
35:34 – 37:23	Fatima generates a new graph and saves it as a final version. This version is represented by Figure 3.

Source: Elaborated by the authors.

Based on the previous tables, we observed the terms were edited more than once for the graph composition. Thus, each learner seemed motivated to provide a coherent and organized graph to their partner, presumably so that their partner would have an understandable illustration to support their writing. Moreover, the authors in their new texts recurrently used the terms that were present in the graph. Therefore, students did

not have to face a blank page when beginning the composing process, an intimidating restriction that many writers have to face (RAIMES, 1998). Such evidence observed during the task indicates that the tool also served to provide an initial prompt, triggering the possibility of a more fluid writing process.

The participant Laisa produced a graph about the TV show Two and a Half Men and forwarded the image (Figure 4) to Fatima.

Figure 4: Graph edited by Laisa, related to a fanfiction about *Two and a Half Men*
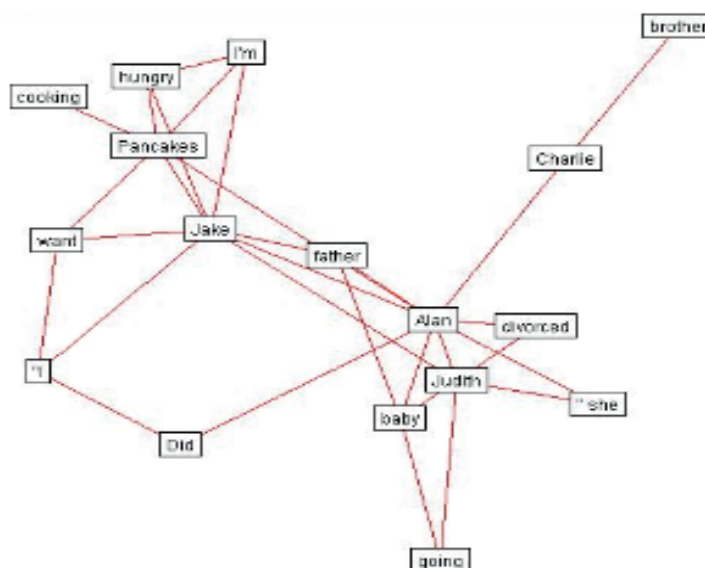
Source: Elaborated by the authors.

Table 3 represents the participant Fatima's process of reading and writing using the *Sobek* tool. A more fine-grained analysis of the text reveals how Fatima capitalizes on the affordances of writing in the fanfiction genre.

Table 3. Fatima's actions when creating her fanfiction

Time	Student's actions
07:16 – 09:00	Fatima observes the graph created by her peer (Figure 4), based on the *fanfiction Mother Again*[6], about the TV show *Two and a Half Men*.
09:01 – 15:17	Fatima starts writing a new story based on the graph.
15:18 – 19:20	Fatima checks the graph again.
19:21 – 27:11	Fatima goes back to writing the fanfiction.
27:12 – 29:17	Fatima checks the graph again.
29:18 – 37:21	Fatima goes back to writing the fanfiction.
37:22 – 38:01	Fatima checks the graph again.
38:02 – 59:32	Fatima goes back to writing the fanfiction and finishes it

Source: Elaborated by the authors.

When writing the fanfiction, Fatima reveals extensive prior knowledge about the TV show. Based on the names of the characters that were in the graph, the student prepared a story according to the scripts of the episodes of the series. She included comic features in the story, the relationship between the brothers (Charlie and Alan) and Jake (Alan's son), conversations between the brothers and Alan's ex-wife (Judith) and even the inclusion of a character that is not in the graph, but is part of the show (Berta). Words included in the graph were used in her fanfiction such as *Jake, Charlie, Alan, Judith, brother, cooking, hungry, pancakes, divorced* and *baby*. Fatima built a strong link with the graph produced by Laisa and the knowledge about the canon, creating an authentic fanfiction related to the original show's scripts.

FINDINGS

After the creation of fanfictions, a questionnaire was delivered to students with the following questions: (1) What do you think was positive and/ or negative about the task you performed? (2) How did you deal with text mining tool (*Sobek*)? How was that experience? (3) When using the tool, how did you build the relations between concepts? What rationale did you use? (4) How do you think the tool can be used in the process of text production? (5) Do you have any other comments? The questionnaire was answered

---

[6] *Fanfiction* retrieved from: <http://www.fanfiction.net/s/5587244/1/Mother-Again>.

in Portuguese, so that students could give their opinion without any language barriers. Therefore, we present a version in English of their answers.

The following Tables show Laisa and Fatima's answers, identified by codes according to Saldaña (2012).

Table 4: Fatima and Laisa's answers to question (1) of the questionnaire

Answers	Codification
Laisa: On the positive side, the text production. I think it was interesting to write a new story from a graph that someone else had created. There was no downside.	Text Production (In Vivo code) Participation of a peer (Descriptive code)
Fatima: I found the task very interesting because it made us exercise our creativity and our skills writing in English.	Creativity (In Vivo code) Text Production (Descriptive code) Practice of the FL (Descriptive code)

Source: Elaborated by the authors.

Table 5: Fatima and Laisa's answers to question (2) of the questionnaire

Answers	Codification
Laisa: The tool is interesting as a way for us to select keywords of a text, and I think it is also a nice way to motivate a creative text production.	Edition of terms (Descriptive code) Creative text production (In Vivo code)
Fatima: It was funny because the graph was kind of strange at the beginning, so I had to remove some words and include new ones. What is cool is that the graph can be created through our selection of words.	Graph (In Vivo code) Edition of Terms (Descriptive code)

Source: Elaborated by the authors.

Table 6: Fatima and Laisa's answers to question (3) of the questionnaire

Answers	Codification
Laisa: I tried to use words that were associated in the text. Then I looked at the graph and returned to the text to see if the words were connected.	Relations in the text (Descriptive code) Rereading (Descriptive code)
Fatima: As I was familiar with the terms and characters, I ended up using my background to make connections. I also thought of what was important in a fictional story (terms like the characters' names, environment, and space).	*Background* (In Vivo code) Narrative elements (Descriptive code)

Source: Elaborated by the authors.

Table 7. Fatima and Laisa's answers to question (4) of the questionnaire

Answers	Codification
Laisa: It (the tool) can be used mainly for those who are learning English, especially beginners I think. But even for more advanced levels, I think it can improve the vocabulary.	Beginners in the FL (In Vivo code) Vocabulary (In Vivo code)
Fatima: It (the tool) suggests some clues in the graph that can support text writing; and it can also stimulate creativity.	Tool as aid for writing (Descriptive code) Creativity (In Vivo code)

Source: Elaborated by the authors.

Table 8: Fatima and Laisa's answers to question (5) of the questionnaire

Answers	Codification
Laisa: It was really cool. I always felt like changing some things about my favorite TV shows or cartoons. At first, I didn't know what fanfiction was and now I'm excited to keep on writing new stories mostly in English.	Interest in the *canon* (Descriptive code) Continuation of authorship (Descriptive code)

Fatima: It was a good experience. I believe it was positive because the graph worked as a facilitator for the writing.	Graph as a facilitator (In Vivo code)

Source: Elaborated by the authors.

The data based on students' perceptions seem to reinforce the screencapture data analyzed in the previous section. Based on the codes identified, we observed that students emphasized the process of text production and the possibilities for creativity enhancement when creating fanfiction. In fact, creativity is a fundamental aspect in the composition of this genre, which was new for some students who got familiar with the process of writing this genre. Thus, the practice of writing in an FL was highlighted by Fatima, who considered the task a valid way to develop facility with the target language. Laisa also stresses the participation of a peer in the brainstorming and preparation for composing, even though the text was later written individually. The process of editing *Sobek*-produced terms in the graphs also yielded some interesting findings about the affordances of both *Sobek* and fanfiction as part of the writing process. Fatima commented on how background knowledge about the canon helped her make connections when creating the graph. In addition, students used the fanfiction they had read originally to constitute relations in the text and edit the terms of *Sobek*. This happened through a process of rereading, which had been already observed in the previous section from the recordings and actions of students. When answering the questionnaire, Laisa mentions revisiting the original text, in order to make sure she was creating coherent connections in the graph. Students viewed the tool as an aid for writing, since there is the visual support of a graph, and also as a resource to work with vocabulary. Therefore, these findings go along with our analysis of the literacy process which was established during the task, since students were able to experience digital and language literacy processes while having to deal with a digital tool and the construction of a new text genre.

DISCUSSION

From the analysis of this reading and writing task, we observed that the graphs generated by the text mining tool and edited by students played an important role in the analysis and subsequent production of fanfictions. The data provided by the participants indicate the graphs helped them reflect and create a clearer image of the stories they

were about to write. The graph, then, was a graphical figure able to reveal a concise representation of a text. The analysis of these images made students go back and forth to the original fanfictions, in order to better understand those texts and create a more complete representation of the stories to their peers. In addition, the writing of fanfictions based on graphs also proved to be a differentiated form of production, based on keywords or key ideas.

According to Sternberg (2006), the fundamental unit of symbolic knowledge is the concept, which can be organized into categories and schemas in our minds. This author states the reader creates a mental model of the text being read, which can be considered a model of inner work of the situation described in the text, as understood by the reader. In other words, the reader creates a mental representation that contains the main elements of the text. Still, according to Sternberg (2006), the construction of mental models illustrates that, in addition to understand them, we also understand how they combine into integrated representations of meaning, when we read a story, for instance.

From this discussion, we conclude that a text mining tool such as *Sobek* can act as a relevant resource to transform this mental model into a visual graph, which acts as a cognitive aid by representing a text through connections between terms. Thus, the tool can provide scaffolding, once students have the opportunity to work with the mining resource while writing authentic texts, such as the genre fanfiction. Hence, the tool may support meaningful writing mediated by technology. Moreover, data suggest that the peer-to-peer exchange of texts was a motivating aspect of the activity. In FL classrooms, it is important to think of activities that involve text production not only targeted at the teacher as reader, who traditionally tends to focus on correction. Activities like the one we did, as well as online environments such as fanfiction writing spaces, can provide authentic audiences for students' communicative practices. Text production, when built in a meaningful way, can turn into the creation of authentic text genres, fostering social practices indeed.

CONCLUSION

This work provides an opportunity to reflect on the possibilities of fostering literacy and language development through tasks of fanfiction creation, supported by the text mining resource Sobek. Thus, we discussed how a digital tool can support learning through pedagogically meaningful tasks. Through this task implementation, students who may never have been part of the fanfiction process have had the possibility of experiencing this

literacy practice for the first time. We consider that such actions mediated by technology can foster students' participation in the process of reading and writing, since they had the opportunity to produce authentic text genres and share them with peers.

As stated by Veen and Vrakking (2007), the generation *Homo Zappiens*, composed especially by people who were born in the late 1980s and later, is characterized by being a new species that acts in a global cyberculture based on multimedia. Previously, Prensky (2001) had already used the term *digital native* to define this generation with individuals who are used to the digital features of the Internet and video games. However, even these connected generations sometimes do not have the opportunity of bringing to classroom their out-of-school digital practices. Learners used to new technologies are sometimes active individual outside school, having control over the tools they manipulate and tending to solve collectively problems that arise while they are connected (VEEN; VRAKKING, 2007). Nevertheless, these students do not always find opportunities to experience new literacies related to digital devices and language development. Above all, it is necessary to include in classroom digital resources grounded on pedagogical proposals in which construction of knowledge is the factor of change. It is important to see digital tools as facilitator devices for building new relationships of meaning. The computer itself does not foster changes, but tasks that promote interaction and knowledge building mediated by digital tools may boost language learning. This way, digital literacy and language literacy can indeed benefit from the use of technology in a contextualized manner.

REFERENCES:

ATWELL, N. *In the middle*: new understanding about writing, reading, and learning. Portsmouth: Cook, 1998.

BAMFORD, A. The Visual Literacy White Paper. *Uxbridge:* Adobe systems incorporated. 2003. Retrieved from: <https://www.aperture.org/wp-content/uploads/2013/05/visual-literacy-wp.pdf>. Access on: 20 jan. 2020.

BARTON, D.; HAMILTON, M. Literacy practices. In: BARTON, D.; HAMILTON, M.; R. IVANIC, R. (Ed.). *Situated literacies*: reading and writing in context. New York: Routledge, 2000. p. 7-15.

BLACK, R. Language, culture, and identity in online fanfiction. *E-Learning and Digital Media*, v. 3, n. 2, p.170-184, 2006.

Black, R. Digital design: English language learners and reader reviews in online fanfiction. In: KNOBEL, M.; LANKSHEAR, C. (Ed.), *A New Literacies Sampler.* New York: Peter Lang, 2007. p. 115-136.

BLACK, R. *Adolescents and online fan fiction*. New York: Peter Lang, 2008.

BLACK, R. Online fanfiction, global identities, and imagination. *Research in the Teaching of English,* v. 43, n. 4, p. 397-425, 2009.

CALKINS, L. *The art of teaching writing*. Portsmouth: Heinemann, 1994.

FARKAS, K.; POLTROCK, E. Online editing mark-up models, and the workspace lives of editors and writers. *IEEE Transactions in professional communication, v.* 38, n. 2, p. 110-117, 1995.

FELDMAN, R.; SANGER, J. *Text Mining Handbook*. Cambridge: Cambridge University Press, 2006.

GRAHAM, S. Strategy instruction and the teaching of writing: a meta-analysis. In: MACARTHUR, C.; GRAHAM, S.; FITZGERALD, J. (Ed.). *Handbook of writing research.* New York: The Guilford Press, 2006. p. 187-207.

GROBELNIK, M.; MLADENIC, D.; JERMOL, M. Exploiting text mining in publishing and education. *Proceedings of the ICML-2002 Workshop on Data Mining Lessons Learned*, p. 34-39, 2002.

HALL, T.; STRANGMAN, N. Graphic Organizers. *National Center on Accessing the General Curriculum*. 2002. Retrieved from: <http://aim.cast.org/sites/aim.cast.org/files/NCACgo.pdf>. Access on: 20 jan. 2020.

JENKINS, H. *Textual poachers*: television, fans and participatory culture. New York: Routledge, 1992.

JOHNSON, D. *Approaches to research in second language learning*. New York: Longman, 1992.

JOHNSON, S. *Interface culture*. New York, NY: Basic Books, 2001.

KERN, R.; WARSCHAUER, M. Theory and practice of network-based language teaching. In: WARSCHAUER, M.; KERN, R. (Ed.). *Network-based language teaching*: concepts and practice. New York: Cambridge University Press, 2000. p. 1-19.

KIM, A.; VAUGHN, S.; WANZEK, J.; WEI, S. Graphic organizers and their effects on the reading comprehension of students with LD: a synthesis of research. *Journal of learning disabilities*, v. 34, n. 2, p. 105-118, 2004.

LANKSHEAR, C.; KNOBEL, M. Sampling "the new" in new literacies. In: M. KNOBEL, M.; LANKSHEAR, C. (Ed.). *A new literacies sampler.* New York, NY: Peter Lang Publishing, 2007. p.1-24.

LEE, L. Blogging: promoting learner autonomy and intercultural competence through study abroad. *Language Learning & Technology*, v. 15, n. 3, p. 87-109, 2011.

LEMKE, J. Metamedia literacy: transforming meanings and media. In: REINKING, D.; MCKENNA, M.; LABBO, L.; KIEFFER, R. (Ed.), *Literacy for the 21st century*: technological transformation in a post-typographic world. New York: Erlbaum, 1997. Retrieved from: <http://academic.brooklyn.cuny.edu/education/jlemke/reinking.htm>. Access on: 20 jan. 2020.

MANOLI, P.; PAPADOPOULOU, M. Graphic organizers as a reading strategy: research findings and issues. *Creative Education,* v. 3, n. 3, p. 348-356, 2012.

MARZANO, R.; PICKERING, D.; POLLOCK, J. *Classroom instruction that works*: research-based strategies for increasing student achievement. Denver: ASCD, 2001.

NOVA, C. Imagem e educação: rastreando possibilidades. In: ALVES, L.; NOVA, C. (Org.). *Educação e Tecnologia*: trilhando caminhos. Salvador: Editora da UNEB, 2003. p. 180-196.

OGATA, H.; YANO, Y.; WAKITA, R. CCML: Exchanging marked-up documents in a networked writing classroom. *Computer Assisted Language Learning*, v.11, n. 2, p. 201-214, 1998.

RAIMES, A. *Exploring through writing*: a process approach to ESL composition. Cambridge: Cambridge University Press, 1998.

PRENSKY, M. Digital natives, digital immigrants. *On the Horizon*, v. 9, n. 5, 2001. Retrieved from: <http://www.marcprensky.com/writing/Prensky%20-%20Digital%20 Natives,%20Digital%20Immigrants%20-%20Part1.pdf>. Access on: 20 jan. 2020.

ROMERO, C.; VENTURA, S. Educational data mining: a survey from 1995 to 2005. *Expert systems with applications*, v. 33, n. 1, p. 135-146, 2007.

SALDAÑA, J. *The coding manual for qualitative researchers*. Los Angeles: Sage. 2012.

SCHWIND, B. Error analysis and explanation in knowledge based language tutoring. *Computer assisted language learning*, v. 8, n. 4, p. 295-324, 1995.

SCRIBNER, S.; COLE, M. *The psychology of literacy*. Cambridge: Harvard University Press, 1981.

SEVINDIK, T.; DEMIRKESER, N. Digital mining applications in distance education environments. *Scientific research and essays*, v. 5, n. 21, p. 3213-3221, 2010.

STERNBERG, R. *Cognitive psychology.* Belmont: Cengage Learning, 2006.

SUZUKI, A.; SATO, T.; AWAZU, S. Graphic display of linguistic information in English as a Foreign Language reading. *TESOL Quarterly*, v. 42, p. 591-616, 2008.

THOMAS, H.; HEALY, A. A comparison of rereading benefits in first and second language reading. *Language Learning,* v. 62, n. 1, p.198-235, 2012.

TOMKINS, G. *Teaching writing*: balancing process and product. Columbus: Merrill Publishing Company, 1990.

TORRANCE, M.; GALBRAITH, D. The processing demands of writing. In: MACARTHUR, C.; GRAHAM, S.; FITZGERALD, J. (Ed.). *Handbook of writing research*. New York: The Guilford Press, 2006. p. 67-80.

VEEN, W.; VRAKKING, B. *Homo Zappiens*: growing up in a digital age. New York: Network Continuum, 2007.

VEKIRI, I. What is the value of graphical displays in learning? *Educational Psychology Review,* v.14, p. 261-312, 2002.

WARSCHAUER, M. Millennialism and media: language, literacy, and technology in the 21st century. *AILA Review,* v. 14, n. 14, p. 49-59, 2001.

WARSCHAUER, M. *Technology and social inclusion: rethinking the digital divide.* Cambridge: MIT Press, 2003.

WILLIS, D.; WILLIS, J. *Doing task-based teaching.* Oxford: Oxford University Press, 2007.