UNIVERSIDADE FEDERAL DE PELOTAS

Centro de Letras e Comunicação

Programa de Pós-Graduação em Letras

Doctoral Dissertation

Automated writing evaluation:

A posthuman perspective on the development of writing in EFL

Gisele Medina Nunes

Pelotas, December 2019

Gisele Medina Nunes

Automated writing evaluation:

A posthuman perspective on the development of writing in EFL

Doctoral dissertation submitted to the Graduate Program in Languages, Concentration Area Acquisition, Variation and Teaching of the Federal University of Pelotas as a partial requirement to obtaining the title of Doctor in Applied Linguistics.

Advisor: Professor Dr. Vilson José Leffa

Pelotas, December 2019.

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

N972a Nunes, Gisele Medina

Automated writing evaluation : a posthuman perspective on the development of writing in EFL / Gisele Medina Nunes ; Vilson José Leffa, orientador. — Pelotas, 2019.

169 f. : il.

Tese (Doutorado) — Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas, 2019.

1. Automated feedback. 2. Foreign language learning. 3. Written communication strategies. 4. Educational software. 5. Posthuman language learning. I. Leffa, Vilson José, orient. II. Título.

CDD : 420.7

Elaborada por Aline Herbstrith Batista CRB: 10/1737

Gisele Medina Nunes

Automated writing evaluation:

A posthuman perspective on the development of writing in EFL

Tese aprovada, como requisito parcial, para obtenção do grau de Doutor em Linguística Aplicada, Programa de Pós-Graduação em Letras, Centro de Letras e Comunicação, Universidade Federal de Pelotas.

Data da Defesa: 11 de dezembro de 2019.

Banca examinadora:

Prof. Dr. Vilson José Leffa
Orientador/Presidente da banca
Universidade Federal de Pelotas

Profa. Dra. Camila Gonçalves dos Santos do Canto
Membro da Banca
Universidade Federal do Pampa

Profa. Dra. Gabriela Bohlmann Duarte
Membro da Banca
Universidade Federal do Pampa

Prof. Dr. Marcus Vinicius Liessen Fontana
Membro da Banca
Universidade de Santa Maria

Prof. Dr. Rafael Vetromille-Castro
Membro da Banca
Universidade Federal de Pelotas

Acknowledgments

Going against the principles I follow in language learning, I will be chronologically linear in thanking some important people who made this study possible.

Thanks to my husband, **Gustavo Stosch**, who blindly supported my decision to be 4 years involved in a life project and willingly gave up so many things so I could make my dream come true. There are no words to thank you enough.

Thanks to **Gabriela Duarte**, who, back in 2015, sat with me for a coffee and walked me through the first steps into outlining a project for applying to this post-graduate program.

Thanks to **Renata Pinheiro**, teacher and friend who helped me build my project. Without your support, I would have never come so far.

My most sincere thank you to my in-laws, **Carlos** and **Laura Stosch**, who made this endeavor financially possible, always reaching out for me. Without you, this moment would not have been possible.

My gratitude to **Elisa Hönnicke**, a colleague and friend who supported me during these 4 years, always having great generosity and sensitivity to deal with me in my most delicate moments.

A kind thanks to my bosses, **Fábio Leite** and **Tiago Nunes**, who also made this journey economically viable, gladly investing in their employee.

A truthful thank you to my advisor, **Vilson Leffa**, who believed in me and pushed me to always go further and trust my potential. Now, I understand what to do research is all about a bit more.

Thanks to **Vanessa Fialho**, whose attention and words helped me kick off the writing of this doctoral dissertation.

A Warm thank you to all my professors and classmates for the timely kind words so needed throughout this process.

Many thanks to my friends, **Gonçalo Cholant** and **Douglas Batista**, for listening to me when things were not easy and specially when no one else was able to comprehend the anxiety and fear we feel when we are post-graduate students.

Thanks to my close friends who understood my several moments of absence, physically and mentally.

A special thanks to my mother and brother, who feel very proud of how far I managed to get even when the odds seemed entirely against me.

Last but not least, thanks to Google, for enabling the search of numerous references at an incredible fast speed. Without this technology, the completion of this study would have been a whole lot harder and longer.

“You’re afraid of change
I don’t know the future
I didn’t come here to tell you
How this is going to end
I came here to tell you
How it’s going to begin
Where we go from there
Is a choice I leave to you”
(The Matrix, 1999)

Abstract

NUNES, Gisele Medina. **Automated writing evaluation**: a posthuman perspective on the development of writing in EFL. 2019. 167 p. Doctoral Dissertation (Doctorate). Programa de Pós-Graduação em Letras. Universidade Federal de Pelotas.

Digital technology has been permeating language learning for nearly three decades. The English Language, in special, has spread worldwide and the attempts to foster its learning through educational software has increased substantially due to the digital information and communication technologies (DICTs). Online applications with automated feedback bring a new perspective in foreign language learning and teaching practices since a different type of dynamics takes place, the human-machine interaction. Learning via digital technologies has been engendering more autonomous practices, in which learners develop language skills assisted by an automated tutoring system. Taking this scenario as a consolidating reality, it is imperative to analyze how this human-machine dynamics in learning takes place, observing how automatic feedback from an automated writing assessor triggers learners' communicative strategies and thus supports more effective learning. Therefore, this study focus on investigating the free automated writing evaluator called *Write and Improve* which assists learners of EFL to develop their writing skills in an independent way as well as the learner-program-teacher triad of interaction. Under a complex, constructivist as well as post-human perspective of language learning, this doctoral dissertation aims at analyzing the application using qualitative methods. The data were collected from texts produced by three subjects who used the online tool to develop their writing skill in beginner and intermediate proficiency levels. Field notes also complemented data analysis, as two of the subjects were part of one of the researcher's class by the time of the data collection. The data showed automated feedback triggers some communication strategies and there was improvement in learners' writing proficiency after a period of three months using the tool. We conclude the automated writing evaluator analyzed in this study can integrate a hybrid instructional environment, where human and automated feedback can intertwine and excel at offering favorable conditions for EFL writing skills' development. From a posthuman perspective, artificial intelligence and human subjectivity are not to be thought as opposing entities but rather as complementary ones, a merge that might contribute to delineate a new pedagogical paradigm for technology-mediated language learning.

Keywords: automated feedback. foreign language learning. written communication strategies. educational software. posthuman language learning.

Resumo

NUNES, Gisele Medina. **Avaliação automática da escrita**: uma perspectiva pós-humana no desenvolvimento da escrita em LE. 2019. 167 f. Tese (Doutorado). Programa de Pós-Graduação em Letras. Universidade Federal de Pelotas.

A tecnologia digital tem permeado o aprendizado de idiomas há quase três décadas. A Língua Inglesa, em especial, difundiu-se pelo mundo e as tentativas de promover seu aprendizado através de aplicativos educacionais aumentaram substancialmente devido às tecnologias digitais de informação e comunicação (TDICs). Aplicativos com feedback automatizado trazem uma nova perspectiva para práticas de ensino e aprendizagem de línguas estrangeiras uma vez que uma dinâmica diferente ocorre – a interação homem-máquina. A aprendizagem via tecnologias digitais tem gerado práticas mais autônomas, nas quais os aprendizes desenvolvem habilidades linguísticas assistidas por aplicativos automatizados. Tendo este cenário como uma realidade em consolidação, é imperativo analisar como essa dinâmica homem-máquina ocorre na aprendizagem, observando como o feedback automático de um avaliador automatizado de escrita aciona as estratégias comunicativas dos aprendizes, proporcionando momentos ricos para a aprendizagem. Portanto, este estudo tem como foco investigar o aplicativo *Write and Improve*, um programa automatizado gratuito projetado para ajudar aprendizes de inglês como LE a desenvolver suas habilidades de escrita de forma independente, bem como a tríade de interação aluno-programa-professor. Sob uma perspectiva complexa, construtivista e pós-humana de aprendizagem de línguas, esta tese de doutorado visa a analisar o programa com base em métodos qualitativos. Os dados foram coletados de textos produzidos por três sujeitos que usaram a ferramenta online para desenvolver sua habilidade escrita com níveis de proficiência iniciante e intermediário. Notas de campo também complementaram a análise de dados, pois dois dos sujeitos faziam parte de uma das turmas da pesquisadora no momento da coleta de dados. Os dados mostraram que o feedback automatizado desencadeia algumas estratégias de comunicação e houve melhora na proficiência escrita dos alunos após um período de três meses usando a ferramenta. Concluímos que o avaliador automatizado de escrita analisado neste estudo pode integrar um ambiente instrucional híbrido, onde o feedback humano e o automatizado podem se mesclar e oferecer condições mais favoráveis ao desenvolvimento da habilidade de escrita em FL. De uma perspectiva pós-humana, a inteligência artificial e a subjetividade humana não devem ser pensadas como entidades opostas, mas como complementares, uma combinação que pode contribuir para delinear um novo paradigma pedagógico para a aprendizagem de línguas mediada pela tecnologia.

Palavras-chave: feedback automatizado. aprendizagem de línguas estrangeiras. estratégias de comunicação escrita. aplicativo educacional. aprendizagem de línguas pós-humana.

List of Figures

List of Tables

List of Appendices

List of abbreviations

CEFR – Common European Framework for Reference

CLC – Cambridge language corpus

CS – communication strategy

DICTs – digital information and communication technologies

EFL – English as a foreign language

FL – foreign language

KR – knowledge of results

LAD – language acquisition device

LSA – latent semantics analysis

ML – mother language

RASP – Robust Accurate Statistical Parsing

TL – target language

UG – universal grammar

VLE – virtual learning environment

ZPD – zone of proximal development

Table of Contents

1 Introduction

We cannot deny the need to master English as a tool to achieve different goals in life nowadays, both personal as well as professional. They can vary from the satisfaction of an international trip without struggle in communication to reaching a disputed position in the job market. Today, knowing English enables people to contribute to the growth of a global society as bilingual speakers (FIGUEIREDO, MARZARI, 2012). In addition, people are living in the era of digital communication, in which being connected to the web is, for many, indispensable. Moreover, technological advances have contributed significantly to the fact that mastering the English language is becoming increasingly dynamic and attractive using enhanced online tools.

With the constant presence of technology in people's daily lives, it seems natural that it plays a more relevant role in language learning. Taking into account the many applications available today, for either computers or mobile devices, such as smartphones and tablets, it is clear that teaching languages using these virtual resources has been gaining visibility. It can be more attractive to this current generation, who grew up using hypertexts, social networks, and video games, learn a language this way since they have already developed a form of reasoning different from that of years ago precisely because of this exposure to technological apparatus (OBLINGER; OBLINGER, 2005).

We can develop four skills when learning a foreign language: speaking, listening, reading and writing. In this last one, to achieve a good performance, it is necessary to master several aspects to produce a good text, such as vocabulary variety and accuracy, correct punctuation, spelling and grammatical rules of the target language, to name a few. In this perspective, positive feedback is extremely valuable to learners' motivation, since it provides emotional support, keeping them active in the learning process (ROSA, 2003; DUARTE, 2017). Likewise, corrective feedback also plays a fundamental role in writing skill development, because this mechanism prevents inappropriate structures to fossilize, for example. Corrective feedback follows certain conventions, such as correction with focus on a specific structure or not, explicit or implicit correction, correction that provides linguistic input to the learner or expects the learner to produce the correct forms (ELLIS, 2009), among other aspects.

Thus, we consider it relevant to observe how feedback takes place on online learning tools since corrective feedback plays an important role in foreign language

learning (FIALHO, 2011). In addition, the format in which feedback is presented to learners can change the way language input is received, causing a series of transformations in the process that can result in varied learning outcomes (CHAPELLE, 2014).

Educational software designed for foreign language writing with automated feedback has been challenging traditional face-to-face teaching, in which a teacher/tutor/instructor is the sole responsible for providing feedback to learners. This scenario is understandable in the sense that it is more economically accessible to a large number of people as well as it is a more comfortable medium for some because it does not necessarily involve human interaction and sometimes delicate judgment issues. Moreover, it can be a mechanism to save substantial teacher time in correcting texts and, consequently, to become a less onerous means of accounting for these written productions in the case of large-scale foreign language proficiency exams.

On the other hand, in light of a complex socio-constructivist view of language learning, the development of proficiency in a foreign language (henceforth FL) is a product of internal and external relations. This means that even if learners engage in well-produced learning activities, online or not, there is no guarantee that it will occur as expected. Besides that, this is accentuated when we observe a more autonomous learning scenario and with an essentially human-machine interaction, which has been spreading at a fast pace on the web and free of charge.

Regarding the context in which the researcher's English teaching practice has been developed, this approach strongly approximates the proposal of the present research concerning the use of technology in the writing skill development. The researcher’s professional pedagogical environment is focused on the search for advancements in FL teaching and learning. These advances are closely linked to the deployment of new classroom technologies that can leverage the process by having the Internet and digital media as great allies. This view arises from the need to find educational practices that include the profile of the generation of digital natives (PRENSKY, 2001), in order to capitalize on the massive time spent connected to the web and achieve more effective results with language teaching.

In relation to the development of writing in FL, the researcher demonstrated in her Master's dissertation positive results with the use of blogs for the improvement of this ability (NUNES, 2013). In her professional field, writing has always been seen as a taboo among learners as well as the most rejected ability in classroom practice. In

view of this local reality but also a common tendency among FL learners in Brazil, influenced by the need for mastering other skills, mainly demanded by the job market (DATA POPULAR INSTITUTE, 2014), we understand the importance of investigating automated applications to develop FL writing skills. Elucidating what they can offer for a positive practice of teachers and students can contribute to redirecting the prestige of writing in FL and perhaps the process of teaching and learning as a whole. We cannot overlook the change generated by the massive use of digital technologies in the most distinct areas of human life and this is a relevant area to focus on.

These new technologies aimed at improving FL writing in an automated way is a field still under development, with little consolidated basis. We understand the need to ascertain the dynamics of these online tools with a deeper and critical look, thinking about how a posthuman language learning and teaching scenario can emerge from this progressive merge between artificial intelligence and human subjectivity. The discussion about the theoretical assumptions that guide these applications is fundamental, especially as regards to their automated feedback, one of the aspects present in the human-machine interaction. Therefore, within this panorama, we present this doctoral dissertation on the potential of learning that these products afford to EFL writing skills development. Automated assistive technologies can excel at promoting the development of writing skills in FL through hybrid work with human subjects when feedback is able to trigger the use of written communication strategies by learners. The social aspect is a determining condition for the individual to acquire knowledge and the new configurations automated writing assessors provide must be investigated in order to delimitate their relevance in complementing the learning process.

1.1. General objective

The general objective of this doctoral dissertation is to observe how EFL learners develop their writing skills using an automated writing evaluation system. We direct our focus to its most distinctive feature – the automated feedback – and the way it triggers learners' written communication strategies in human-machine interaction. To account for this endeavor, we selected the British application *Write and Improve* to provide this investigation with the automation element. Observing this dynamics, we aim to identify how this type of technology can better assist human work on feedback

and compose an instructional scenario that incorporates the best of both worlds - digital and human - in order to achieve more satisfactory pedagogical results.

1.2. Specific objectives

Our specific objectives focus on

a) observing the types of errors that the program identified in the subjects' texts;

b) detecting the form of presentation of the feedback to the subjects;

c) analyzing the triad of interaction that the interface enables through the +*Class View* utility, involving learners, application and teacher; and finally

d) examining the written communication strategies that the automated feedback incited in learners so that they could solve their communication gaps and reach a higher performance level.

In order to address the objectives presented, we judged the qualitative methodology as the most favorable way, given the documental and exploratory nature of this investigation. In view of this panorama, the objective here proposed is not to find a finite answer to a given question, but to characterize a process in constant transformation, once we consider language learning a highly adaptive phenomenon that takes unpredictable directions.

1.3. Dissertation structure

This doctoral dissertation is divided into 8 chapters. In the first chapter, we present the introduction, mentioning the researcher's background, followed by the current scenario in teaching and learning mediated by online tools as well as the motivation that gave rise to this study. A the end, we bring the objectives set for this investigation and the dissertation structure.

The second chapter brings a review of the literature regarding automated writing evaluators and their improvement throughout the last few decades. We also present studies concerning their efficacy when applied in different contexts of instruction.

In chapter 3, we account for the feedback and its essential role in language learning. We show its different taxonomies, sources, objectives, and transformations in relation to the main second language acquisition paradigms. At the end of this

chapter, we focus on automatic feedback and its connection to written communication strategies. We present studies about written communication strategies, showing the relevance of their use in the improvement of learners' writing skill. We bring the description of some of the most common strategies employed by students when trying to convey their thoughts.

In chapter 4, we introduce the terms interaction and interactivity, its differences and meanings in different areas of knowledge. We also discuss interactivity as playing an essential role in providing means for learners to establish a communication channel with the tool and benefit from an efficient human-machine interaction.

Chapter 5 contemplates the methodology we employed in this investigation. We present the research context in which the experiment took place in detail, the subjects who participated in the experiment that generated part of the data we analyzed, the data collection methods and, finally, the analysis procedures we employed to reach the results intended.

In chapter 6, we bring the data analysis and discussion of the findings, starting from the data collected by the researcher herself to document the structure and functionalities available on the online tool. Next, we look into the data generated by the participants observing how the human-machine interactions took place when an automated pedagogical tool supports learning. We analyze the types of feedback available having the learners as the ones who initiate interaction dynamics and how written communication strategies were triggered off by the application feedback.

In chapter 7, we intend to present the challenges found in working with such a tool, always under the perspective of how human-machine relations engender an extremely lively and oscillating process as language learning is. We will point out the flaws we found in regards to assisting learners in reaching out to the appropriate communication strategies. The same way, we will present the possibility of learners' improvement when practicing writing autonomously with *Write and Improve*. At the end of this chapter, we plan to suggest some directions that automated writing assessors should take in order to assist independent learners to profit better from this new learning environment.

Chapter 8 contains the references used to guide this study followed by the appendices section.

2 Automated writing evaluation systems

A considerable number of studies about automatic feedback in the area of language learning and teaching focus on the improvement of the written ability. With the use of specific software, the students write their texts and receive automatic feedback of their errors, usually according to a database previously established in the program. The most common definition of these programs is a computational technology developed with the purpose of evaluating and punctuating written prose (SHERMIS; BARRERA, 2002; SHERMIS; BURSTEIN, 2003; SHERMIS; RAYMAT; BARRERA, 2003). They can assist teachers both in the context of classroom assessment as well as companies that apply large-scale proficiency tests such as IELTS, TOEFL, TOEIC, PTE and English Language proficiency tests at the University of Cambridge, England, to mention the most notable.

Although there is still no consensus in terms of taxonomy, these programs have become visible in the literature of the area as AWE, the acronym for Automated Writing Evaluator. Other less common denominations may be found for such technology such as Automated Writing Feedback or Intelligent Tutoring Systems to name a few. It is important to emphasize that some authors prefer the denomination that brings the idea of evaluation (AWE), once they believe in its wider scope, including not only score but also the question of a wider assessment with other forms of feedback and tools available to assist writing (CHEN; CHENG, 2008).

The first program developed for this purpose was the Project Essay Grading (PEG) in 1966 by Ellis Page, to make large-scale evaluation more practical and effective (RUDNER; GAGNE, 2001; PAGE, 2003). Although always intended to resemble correction from a teacher, the system at that time was restricted in technology for such, mainly because it ignored semantic questions in the evaluation and organizational and style aspects of a text. The lack of these aspects made the system easy to trick and consequently decreased its credibility. In addition, PEG did not offer feedback to learners, which was also considered a limitation. The system was upgraded in the 1990s in metrics, parsers, and classification schemes, but today it is no longer an available resource. The following illustration was the clearest image found of PEG on the search engine Google:

Figure 1 - Project Essay Assessor's interface

In 1998, another software with this purpose was released, called Intelligent Essay Assessor (IEA), by professors Thomas K. Landauer, Darrell Laham, and Peter Foltz, of the University of Colorado, United States. The researchers tried to reduce costs and speed the process of text evaluation with a technology that sought the closest possible resemblance to a human evaluator. In news about the tool release, Lahan commented, "in a way, it tries to mimic the function of the human brain." Thus, the tool is able to evaluate not only expository texts but also creative narratives, offering feedback to the learner and checking the text against plagiarism.

In contrast to other systems, its focus is to point out inconsistencies in the quality of text content. However, it also spots inadequacies in formal aspects of the language such as grammar, spelling, and punctuation (STREETER; PSOTKA; LAHAM; MACCUISH, 2004). To evaluate the texts, the technology used is the Latent Semantics Analysis (LSA), which, in general, consists of transforming the semantic content of words of a text into vectors, by means of a mathematical calculation of angles, creating a space of semantic representation of the information contained in the textual composition. Vectors of different texts are compared and their quality is established in terms of the content conveyed, originating a score.

The IEA developer, Pearson Knowledge Technologies, claims that the tool is not able to evaluate more intrinsic aspects such as individual creativity and critical

thinking present in a text (DIKLI, 2006). The following is an illustration of how the users saw the evaluation of their text, as shown in figure 1:

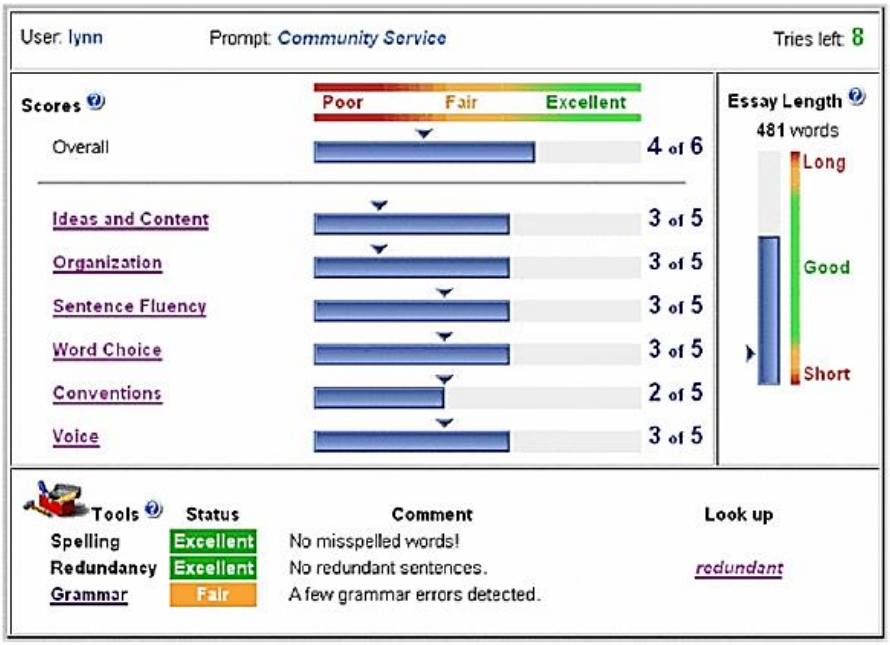


Figure 2 - Intelligent Essay Assessor's interface

At the top of the screen, there is the overall evaluation of the submitted text indicated by colors in a continuous line. On the right side, there is text size evaluation also based on a solid line, where green represents a good size of text and the ends of the line with yellow and red colors represent inadequate sizes, too long or too short text. In the center of the screen, there are more detailed elements that have been evaluated, such as ideas and content presented, organization of the text, fluency of sentences, choice of vocabulary, conventions (such as ABNT and APA) and voice (covering style, choice of words and punctuation). At the bottom of the screen, there are aspects such as spelling, grammar, and redundancy, with evaluative expressions (excellent and good) and superficial feedback on these elements. As this tool is no longer available, from image 1 it is impossible to check whether the learner could visualize their errors in the text and how they were presented in visual terms.

Also at the end of the '90s, the North American Educational Testing Services (ETS) developed a program called e-rater, with the purpose of evaluating the linguistic quality of texts written in English. Its first application was the evaluation of the texts from the Graduate Management Admissions Test Analytical Writing Assessment

(GMAT AWA), an exam for admission to postgraduate programs in the area of administration, such as the Masters of Business Administration (MBA). However, since February 2006, the AWA GMAT exams have been evaluated by another system, called Intellimetric, which was developed by Vantage Learning (ATTALI; BURSTEIN, 2005; DIKLI, 2006).

Both e-rater previously and Intellimetric these days had as purpose the replacement of human evaluators. Two people were in charge of evaluating each text and, if there was disagreement among them, a third evaluator was supposed to conclude the process. Today, only one person and the automated system perform the assessments and, if there is a discrepancy in the results, a second human appraiser is included in the process. In this way, the system's ability to perform the task of evaluating written text on the same level as a human assessor is evident, supplying a probable market need to make this business follow more economically sustainable standards.

The technology behind e-rater and Intelimetric is the natural language processing, one of the areas of artificial intelligence, which focuses on programming systems for understanding and manipulating natural human language. It has been applied to tutoring systems since the 1990s and has been used to train them with large amounts of texts previously codified by human analysts, becoming able to understand the rules that underlie the linguistic code and consequently detect inconsistencies in language use.

Currently, e-rater is part of Criterion's interface, also a product by ETS, with the same purpose of evaluating texts written by English language learners. In the company webpage, the service is presented with the message "more teaching, less correcting", which implies the idea of no longer spending so much time correcting students' texts and investing more in teaching time. Its proposal includes helping the students plan, write and revise their essays, receiving immediate feedback – the automated one. However, a video demonstration of the company mentions that the texts feedback takes about ten minutes to reach the user. This feedback acts as a complement to the summative assessment. It approaches aspects such as a) organization and development, offering feedback on transition words and progression of ideas in the text; b) grammar, which includes spelling, verbal inflection, and agreement; c) language use, which deals with the use of prepositions, articles, word classes and non-standard constructions; d) mechanics, which addresses errors related to punctuation,

hyphenation and use of the apostrophe, and d) style, which provides feedback on very repeated words, very short or very long sentences, and passive voice. The following is an image of the Criterion interface, viewed by the user.

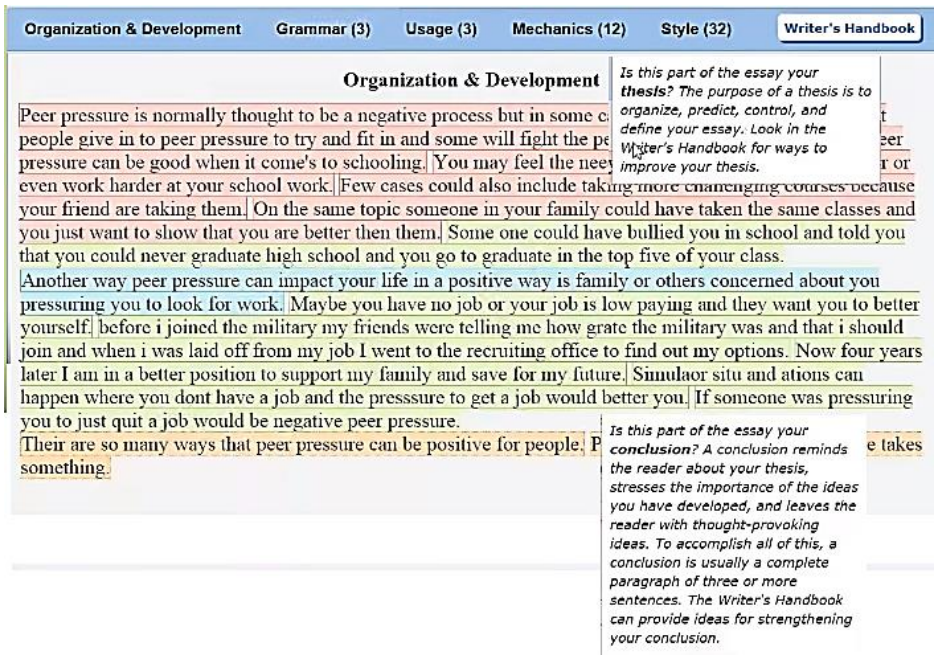


Figure 3 - Criterion's interface

Over the years, Criterion has expanded to cover various levels of learners, ranging from the elementary level to the higher education level. It is also able to evaluate texts of EFL learners. With regard to the texts that it evaluates, the program can work with several genres and, according to its creators, it requires only 465 texts of a certain topic previously corrected by specialists to be trained and perform the assessment. One of the most recent enhancements to the program is the teacher being able to include feedback comments to the texts that their students submitted to Criterion. Nonetheless, it is important to emphasize that this system is not freely available and American educational institutions use it to complement educational practice. Thus, institutions are responsible for affording it so teachers can use in their classes, making the writing practice more agile and with more technological support to the teacher.

Regarding the effectiveness of Criterion, a study by Li et al (2015) presents the reaction of students and teachers to its implementation in the classroom in the form of supporting technology. The system was used to complement writing classes in English

in order to free the teacher from correcting grammatical, spelling and punctuation errors so that he could focus on content issues, ideas and the organization of the text as a whole. In general, students and teachers approved the technological support offered by the program but pointed out that some corrections seemed incoherent and confusing. In addition, although the program also provides feedback regarding the organization of the text, both students and teachers considered it superficial. The correction was based on a pre-established textual standard, which limited it to provide feedback interpreting questions of individual style and creativity in the productions of more advanced learners.

My Access is another paid online writing assessment tool that has the Intellimetric system technology as support, which uses natural language processing for text analysis, both products of the American company Vantage Learning, created in 1998. On its official website, the company presents itself as the world leader in educational technology, online assessments, and instructional programs, and claims to be the first institution to use artificial intelligence to offer the automatic text evaluation service. My Access mainly invests in learners up to 12 years old and its campaign on the website advocates the power of the tool to improve significantly the ability to write, through the presentation of some statistics and comments from teachers. There are also versions for practicing writing at university level (My Access College Prep) and for preparing for the GMAT AWA test (My Access Professional Edition), both with a lower content of advertising.

The main objective of the program is to offer its users a virtual environment in which they can practice their writing skill and receive automatic feedback on their compositions. Currently, the system enables the practice of writing in English, Spanish and Chinese, with the intent of expanding the range of languages available in the future. It requires only 300 texts previously corrected by specialists to be trained and perform the assessment (VANTAGE LEARNING, s.d.). One of its differentials is that the learners can choose the language they want to receive the feedback in among the three available in the program. Another purpose of this tool is to provide more time for teachers and parents, in the case of children receiving instruction at home, to dedicate themselves to teaching and to spend less time correcting the written tasks of their students and/or children, since it is able to examine various textual genres. Figure 4 shows My Access' interface, retrieved from Mohsen and Alshahrani (2019, p. 124).

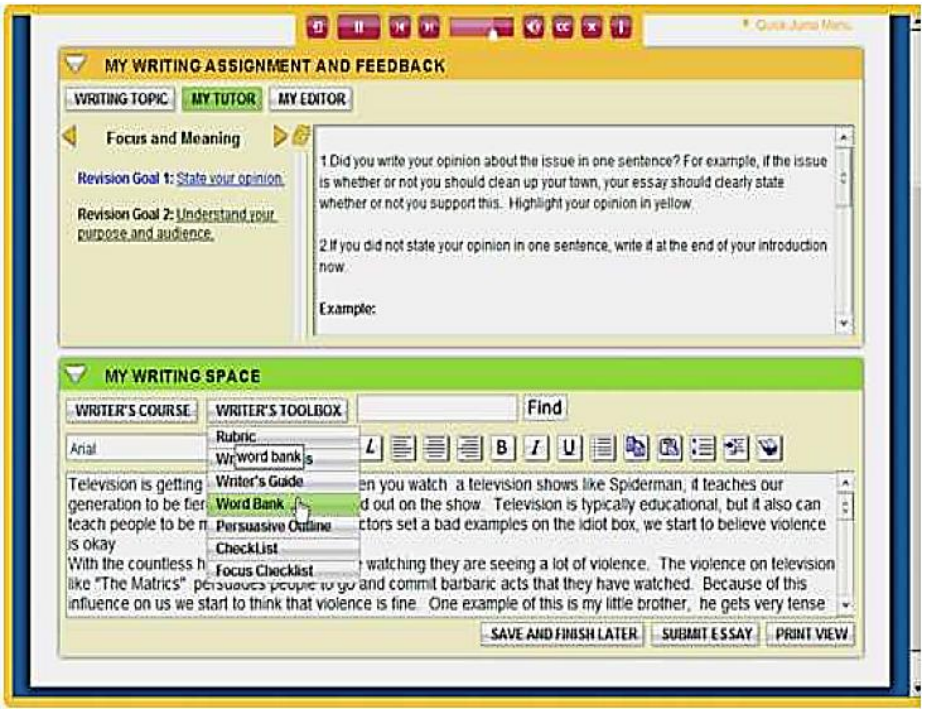


Figure 4 - My Access' interface

The authors proposed in their study a hybrid mode for correcting students' essays in EFL. Along with the automated feedback provided by the tool, there was a teacher in charge of the experiment who provided complementary oral feedback. The system gives holistic grades to the texts from 1 to 6. All submitted texts had higher scores comparing to their first drafts, implying the automated feedback was helpful. When the authors compared the writing skill improvement using the hybrid mode of feedback and only the automated feedback provided by My Access, they concluded that making use of both artificial and human feedback yielded better results in subjects' writings. This suggests that developing strategies that include the automaticity of an online writing tutor intertwined with feedback coming from a teacher can produce successful results in FL writing skills development.

In regards to the effectiveness of My Access to improve writing skills, Yang (2004) showed a smaller acceptance of this type of tool for the correction of texts by more proficient students. However, once verifying Vantage Learning research reports concerning its product, the available documents show that the same study pointed out that most students had a positive reaction to the tool and found it easy to use. Such an assertion is unreliable to the results of the study in question, resulting in an inconclusive reckoning about the real capacity of the program.

The study by Chen and Cheng (2008) investigated My Access College Prep version with Taiwanese students at university level in the third year of the English language course with proficiency level considered intermediate. The authors' work sought to determine the effectiveness of the tool in the context of higher education classroom as well as the implications of implementing this type of technology in an academic environment. The authors analyzed three different groups, ranging from 18 to 26 members, conducted by three teachers with experience in EFL writing and following a similar teaching program to develop academic writing. The total number of classes was 18 weeks, with weekly meetings lasting three hours each. In order to include the online program in their classes, the teachers had an hour-long theoretical instruction with a Vantage Learning consultant. For the practical aspects of the system, teachers needed to work on their own in order to learn more about its functionalities before implementation. Besides, none of the students had any experience with automatic writing assessment programs.

Using the analysis of the program feedback to submitted texts, questionnaires and interviews with teachers and their students, the study concluded that from the three teachers surveyed, only one relied on the evaluation of the program to correct the texts of the students. The other two participating teachers mentioned aspects of the software limitation in the sense that it restricted the aspect of the topic of the text, its structural organization, and the discursive style. In this way, the students' creativity was compromised by the restrictive format of the program, an aspect that was not considered useful for their writing development. In addition, one teacher also commented on the fact that he often had his work doubled by the use of the program, since there were times when he needed to explain and/or revise the vague corrections made by the tool, which became more extenuating than if he had only applied his own feedback on the assignments.

The research by Chou et al (2016) also evaluated the effectiveness of My Access in improving the written ability of English language learners. The study analyzed five pre-intermediate learners studying Administration, Chinese Language, and Accounting. As in the previously mentioned study, My Access worked as an auxiliary tool to these students' writing classes. The purpose was to submit as many revisions of a narrative as possible for the period of two months. The study concluded that some participants found it difficult to improve the construction of their narrative because they did not understand the feedback once their level of proficiency was not

compatible with the complexity of the feedback provided by the system. In addition, the feedback was always very generic and decontextualized, causing frustration and doubt about what students should modify in the compositions. On the other hand, the participants considered the tool useful when they needed to submit a large number of assignments, because of the constant revisions.

Outside the commercial focus, the Bayesian Essay Test Scoring System (BETSY), developed by Lawrence Rudner, is considered a research tool and not a product itself and it was designed exclusively for Windows (BETSY, s.d). It is based on the Bayesian theorem, which includes two text evaluation models: "the Bernoulli Model checks for the presence or absence of a given linguistic aspect in each text and the Multimodal Model confers the multiple uses of a given linguistic aspect in a text" (RUDNER; LIANG, 2002, p. 7). Its focus is on small texts and the technology used is the same as other systems mentioned before: BETSY needs training with texts previously analyzed to build its database. The tool needs 1000 texts to learn how to perform the assessment of texts.

According to Dikli (2006), BETSY is considered an evolution of other systems allied to its own characteristics, giving it a unique format. However, the training is performed manually, and who wants to use BETSY needs to feed it with the texts of his/her own choice. It is necessary to create projects, which delimitates the type of analysis desired (Bernoulli or Multimodal), determine whether the analysis will be word-level, phrase or argument, as well as several other items. In this way, the users will customize the program to examine specific aspects of their preference in the texts they submit. It is pertinent to point out that for FL learners who intend to develop their writing and nonprofessionals in the field of computing and its respective language, BETSY is extremely complex to handle. There is a need for a broad familiarity with the program in order to understand its commands and execute them appropriately to achieve the desired results. Given its peculiarities, the program seems to be inappropriate for the public. On figure 4, a screenshot of a test made by the researcher on Betsy to illustrate its interface:

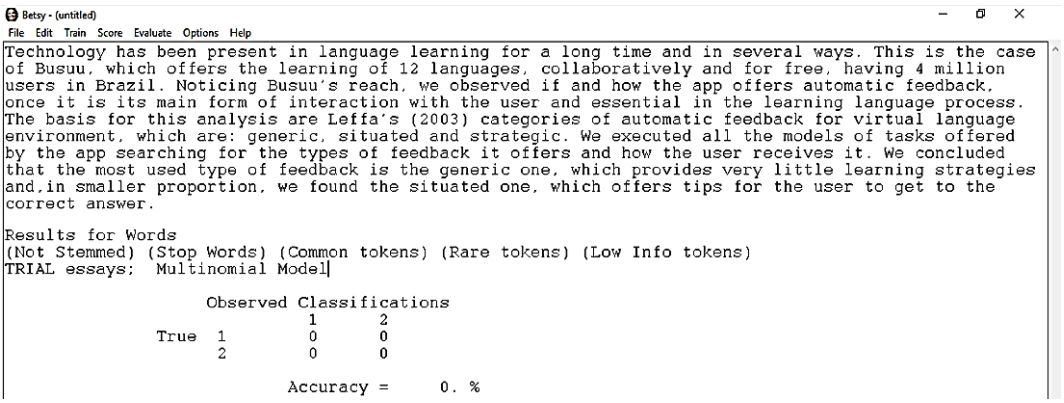


Figure 5 - Betsy's interface

Another paid program for the assessment of texts in English is Writing-Pal, considered an intelligent tutoring system (MACNAMARA et al, 2012), developed with the focus on offering strategies to improve the writing of students of educational institutions at the high school and higher education levels. The tool is based on three stages for learning writing: planning, writing and reviewing. The automatic score is based on natural language processing algorithms. The feedback brings animated videos, which instruct the learners on what to do at a particular stage of writing, using examples, reference guides, and mnemonic techniques. The developers consider the feedback individualized, formative and gamified. In addition, there are interactive assessments in the format of questionnaires and games included in the process. Writing is done by choosing a template from the tool that is available and a time limit to complete the task. Feedback, said to be formative, aims to assist the learners in identifying and selecting appropriate strategies to improve their text. According to Macnamara et al (2012, p.500), the automatic response to the learner can be informed from a more computational or pedagogical perspective. If the algorithm included the number of words aspect and it is interpreted in the light of a pedagogical objective, the feedback would be "elaborate more ideas" and not "add more words", which illustrate what the author means by constructive feedback. Following, an illustration of My Writing Pal's interface (Figure 6).

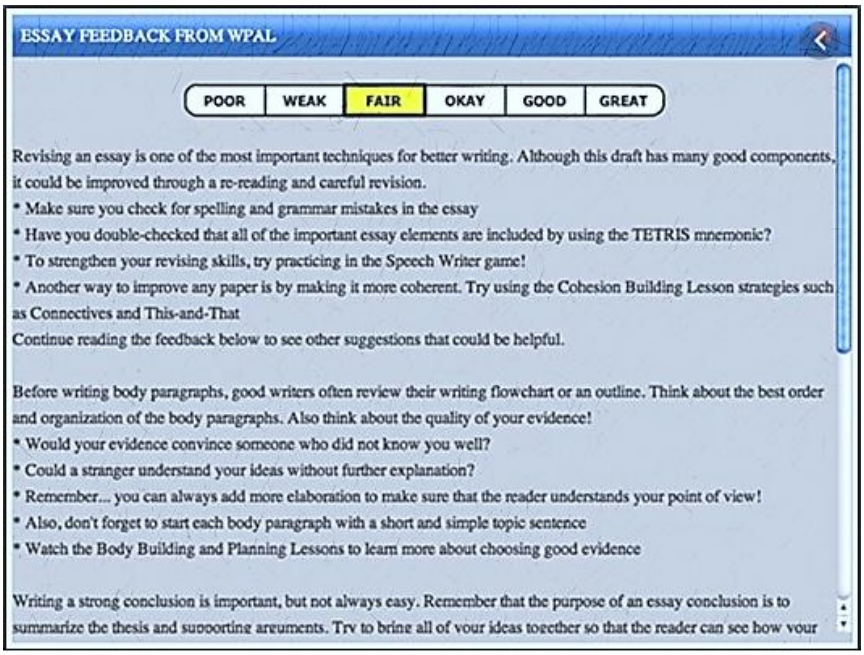


Figure 6 - My Writing Pal's interface

In a study with the tool implemented as a support to the lessons, Roscoe and Macnamara (2013) investigated its incorporation and effectiveness in a US school with elementary students over a period of 6 months. The teachers had weekly meetings with the researchers and could clarify any doubts or needs regarding Writing Pal functionalities. The average system usage per student was 16 hours. The results showed that in terms of access frequency, its use decreased progressively over the period. In the first three months, there was intensified use and, by the fifth month, the use was very sporadic. Such repercussions are justified by the decreasing incentive on the part of teachers to manipulate the system. Thus, since the teacher did not impose it, the students showed a growing disinterest in using the program. Regarding Writing Pal students' perceptions, most of them considered it an easy-to-use program with sufficient feedback. However, a portion of respondents commented in the interview that the feedback could be more detailed, focusing on what exactly should be improved. Finally, the study concluded there was progress in improving writing skills with the help of the tool, once students produced longer, more organized texts with more diversified vocabulary at the end of the observation period.

For a panoramic view of the most relevant systems developed for the automated assessment of English language learners texts and their main characteristics, see table 1:

Table 1 - Comparison of automatic writing assessment systems adapted from Dikli (2006, p.23)

Automatic writing assessing tool	Developer	Technique	Main Focus	Instructional Application	N. of texts needed for training
Project Essay Grade (PEG)	Page (1966)	Statistics	Style	-	100 to 400
Inteligent Essay Assessor (IEA)	Landauer et al (1997)	Latent Semantic Analysis	Content	-	100 to 300
E-rater	Educational Technology Service (ETS) (Burstein et al, 1998)	Natural language processing	Style and Content	Criterion	465
Intellimetric	Vantage Learning (Elliot et al, 1998)	Natural language processing	Style and Content	My Access	300
Bayesian Essay Tutoring System (BETSY)	Rudner (2002)	Bayesian Essay Test Scoring System	Style and Content	-	1000
Coh-Metrix	Macnamara et al (2012)	Natural language processing	Content	Writing Pal	120

According to the aforementioned studies involving these online tools, we can observe that the implementation of an automatic writing assessor in a formal environment, such as schools and universities, requires planning and training the teachers involved. Without minimally accounting for these aspects, there is a high chance of reducing the potential effectiveness of the tool and substantially compromising the teaching-learning process linked to it, making it a totally expendable investment. Having such a technological apparatus in the classroom can be a transformative practice for both teachers and students, considering the necessary engagement of all participants as well as their incorporation in a strategic way.

To support this idea, Leffa et al (2017), in a task-driven study on a Moodle course of instrumental English, investigated if students preferred either individual activities with automatic feedback, individual writing or group discussions. The results showed that the students preferred the activities with an automatic response instead of the other two varieties. These data suggest that the question of instantaneity is positive because it can enhance learning since there is no waiting for feedback. Likewise, the students are able to redirect their focus while still performing the task. On

the other hand, the authors argued that the activity with this feedback model – the quiz – did not imply written production by the students, which happened in the other activity formats, which may have influenced their preference, considering the cognitive and affective demand that the task of writing involves (LEFFA et al 2017, p 126). It is worth noting that automated feedback was present only in the quiz activity. The other activities were corrected later by the teacher in charge of the online course. We can suggest that if written production tasks were supported by automatic feedback, students might have changed their view of them.

Moving to smaller educational ventures in the area of automatic writing assessment, Milton (2004) and Milton and Cheng (2010) developed the software Mark My Words on an experimental basis to help language teachers correct their students' assignments at the university level in a more dynamic way. The software is an extension of MS Word and offers three tools to the teacher. The first is a database of 100 most recurring lexical-grammatical errors previously recorded in the program that can be inserted into the text with just a few clicks. In addition to this list of errors, the program contains a set of small predefined metalinguistic comments that the teacher can use to increase their correction. Finally, the teacher can also add links to the text with sources for the student to access more information about their errors and the correct forms, such as a dictionary or an online grammar reference. According to the authors, this strategy aims to induce the students to discover and learn for themselves, developing their autonomy and ability to solve problems without constant direct instruction from the teacher. As it can be seen, comparing the other programs presented previously to the Mark My Words tool, this one is noticeably more limited in relation to the automated feedback that it provides to the learner. In addition, the teacher is constantly involved in the process, having to read the texts to manually activate the program's database. From this perspective, this tool would not be configured as fully automated, because the program does not provide feedback independently. Finally, this extension is currently no longer available for use.

It suffices to say that Milton (2004) uses the term electronic feedback and not automatic or automated feedback as other authors do to label the automatic response that a program sends to a written text to provide feedback. This vocabulary choice is also used by Warschauer (2006) to conceptualize the correction made in the writing of L2 learners mediated by technology. However, Warschauer brings two meanings to dealing with what is electronic. The first one refers, according to him, to the source

from which this feedback comes, that is, from software on a computer, according to Cardoso's (2011) interpretation, in which feedback can be characterized according to its content, direction, moment or source. The second meaning concerns the medium through which feedback spreads. Feedback is electronic because it is mediated by a machine, but is essentially provided by an individual at the other end of the interaction, which does not characterize the feedback of the automatic type, which we are describing in this section.

However, the author addresses an interesting question beyond this clarification of terms: the view that one has of writing that is under the effect of feedback. For him, for teachers who see the domain of writing "mostly as a compendium of sub-skills", to make use of electronic feedback in its automatic sense satisfies the needs (WARSCHAUER, 2006, p.3). In this context, the primary concern is only the form of the text, how precise the structure of the language is within the sentences and the sentences within the canonical five paragraphs of academic writing. On the other hand, for those who see writing as a social practice, a communicative act which requires interaction, feedback needs to be provided to anyone interested in reading and commenting on what was written. In this way, technology serves not as a mere evaluating service, but as a tool that provides access to other written genres as well as learning through collaboration among colleagues, enriching the learning process in both material and human terms. From this constructivist perspective, writing is more than practicing a skill; it is in fact communicating. Thus, writing to software with previously recorded mechanisms may not be such a fruitful strategy precisely because of the human component's gap in reading and interpretation. This aspect is intrinsically linked to human subjectivity, capable of contributing much more than appropriate verb patterns and proper spelling. The reader is able to capture an intention of what the writer wanted to convey and from there we weave feedback directing his/her adjustments to what he/she imagines he/she has understood. In a rudimentary way, it can be said that software detects linguistic elements positioned according to patterns pre-established by the language system, which sometimes leaves out cultural, subjective and social issues of who produced the text. To support this view, Chen and Cheng (2008, pp. 94-95) argue that

> Voices from the academic community presented in Ericsson and Haswell's (2006) anthology, for example, question the truth of the industry's publicity for AWE products and the consequences of the implementation of AWE in writing classes. They distrust the ability of computers to "read" texts and evaluate the quality of writing because computers are unable to understand meaning in the

> way humans do. They also doubt the value of writing to a machine rather than to a real audience, since no genuine, meaningful communication is likely to be carried out between the writer and the machine. Moreover, they worry whether AWE will lead students to focus only on surface features and formulaic patterns without giving sufficient attention to meaning in writing their essays.

As we can see, one of the initial motivations for developing software capable of providing automatic feedback was to substantially reduce the hard work of corrective work by teachers, who are responsible for reading and marking inadequacies in significant quantities of their students' texts. In addition, another key reason was control of the costs involved, considering the time required for massive amounts of texts to be corrected by professionals in proficiency tests applied on an international scale, for example. In order to ratify this statement, Warschauer and Ware (2006, p.6) point out that

> Teams of graders must be given interrater reliability training, and grading must be regulated with reliability checks. These necessary protocols are costly, which makes automated essay evaluation appear as an attractive, economically viable, alternative because replacing human raters with automated raters reduces the overall cost of the evaluation.

According to Ted Briscoe (2015)[1], a computational linguist, educational technology has received heavy investments in the last decade and has been improving exponentially. However, the areas that are most successful with feedback automation are those from the so-called hard sciences. The linguist argues that the MOOC Coursera, for example, is considered a success in this matter, since feedback is always precise for most of the courses offered. There are no divergences, ambiguities or a wide range of answers. And this is what language learning does not have: a single right answer. Language is dynamic, contextual, open to possibilities, not offering this right or wrong duality. Therefore, automatic feedback is not so effective for the learner if the program is not able to deal with the nuances the language holds in certain cases. There seems to be agreement among the experts in the field regarding the format that feedback should have in order to be more effective in affording learning moments. However, what is observed in language learning mediated by large-scale technology, in practice, is still not what theorists advocate. In general, researchers argue that by using a language learning tool, students can construct this knowledge in an autonomous way, without getting all the answers immediately, being encouraged to

[1] Talk delivered at the eLex 2015 conference. Video available at: https://elex.link/elex2015/videos/

think and look for what they do not know yet. Then, it is preached that the feedback produced by these tools should be formative and not summative. However, what the available applications, free or not, are able to offer so far is still below that demand. What these programs address in terms of autonomy would be, superficially, the fact that the individual uses them without the assistance of a teacher or tutor or undergoes a program of studies established by others.

Research in the area has not been clear in order to prove the positive results in the writing of EFL learners. This circumstance is given because there is not a solid paradigm and consequent appropriate methodology to account for the use of programs for review and correction of writing automatically. Another significant reason is the fact that the determining factor for satisfactory performance in learning writing with the use of these technologies is the pedagogy employed to work with them. The automatic writing assessors cited as examples are inserted in an institutional context, where there is a curriculum, inside a teaching institution and under the responsibility of a teacher to plan and guide the activities. Thus, these aspects are crucial to evaluate the effects that the implementation of this type of technology in the teaching-learning process can cause.

We can suggest, in a first analysis, that developing software with this feedback complexity in terms of managing and delivering information to the user automatically requires a huge investment with programmers, designers, linguists, and teachers, to mention some of the professionals which a project with this specificity implies. Besides, there are costs with infrastructure and many other aspects. Perhaps the return in financial terms to a venture of this magnitude does not cover the amount that is employed and so it is not an attractive business in the eyes of companies that work with technology in education., Although there are people and companies working to improve education constantly, we must keep in mind that we live in a capitalist world and in education, it is not different - profit is an ever-present aspect.

This is undoubtedly a recent field of investigation, but it is evident in the few studies already published that a human part of the feedback work is notably reclaimed when using technology to speed it up and/or automate it. Only the work of software does not fulfill learners' needs in their process of becoming proficient in FL. This type of technology can act as an excellent complement to good teaching pedagogy, but it is not, at least, capable of replacing a teacher completely. In the words of Bates (2015),

> [...] it is likely to be a major mistake to use computers to replace or substitute humans in the educational process, given the need to create and interpret meaning when using media, at least until computers have much greater facility to recognize, understand and apply semantics, value systems, and organizational features, which are all important components of 'reading' different media. But at the same time, it is equally a mistake to rely only on the symbol systems, cultural values and organizational structures of classroom teaching as the means of judging the effectiveness or appropriateness of the Internet as an educational medium.

Therefore, there is a need to better understand the extent of the aid that automated teaching tools can offer and thus compose an instructional scenario that incorporates the best of both worlds - digital and human - in order to achieve more satisfactory pedagogical results.

3 Feedback

3.1 Term origins

The term feedback has its origins in the hard sciences. In 1909 physicist and Nobel Prize winner Karl Ferdinand Braun used the term as a noun for the first time to refer to the connection between components of an electric circuit (Braun, 1909). Robert de Beaugrande also corroborates this origin of the term when quoted in Paiva (2003).

According to the first concept of the American online dictionary Merriam-Webster for the term, it is called feedback

> the return to the input of a part of the output of a machine, system, or process (as for producing changes in an electronic circuit that improve performance or in an automatic control device that provide self-corrective action).

Ramaprasad (1983) defines the term broadly as "information on the difference between the current level and the reference level of a parameter system, which is used to change that difference in some way." In nature, Briggs (1992) points out that feedback instigates the environment to transform itself. It is a survival mechanism of organisms, which goes from the cellular level to ecosystems. Rivonlucri (1994) defines the term feedback in this same line as the message that returns to the body after its action in the environment.

We can notice that from engineering to biology, the concept of feedback is essential to describe the operation of the most varied systems. An example of the engineering field is the thermostat. The device measures the place temperature, compares it with the desired temperature and, using feedback from the difference between the two measurements, turns on to increase the temperature or turns off to adjust to the desired level. An example that illustrates feedback in biology is the regulation of glucose in the bloodstream. By secreting insulin and glucagon through the pancreas, the body tries to keep its glucose level constant. When the glucose concentration rises, the hormone insulin is released, causing the body to store excess glucose in the liver. In the opposite situation, when the glucose level is low, the pancreas secretes the hormone glucagon. This constant dynamics between the liver system and the pancreas system keeps the organism in balance (ÅSTRÖM; MURRAY, 2008).

3.2. Feedback and language learning

In the context of teaching and learning a language, feedback is extremely valuable. It makes the learners adjust their efforts and decide which goals to pursue and which to set aside, depending on the occasion. For learners' motivation, feedback plays an essential role as it provides emotional support, keeping them active in the learning process (White, 2003; Ellis, 2009). Here is a very simple example: the teacher asks a written assignment to his/her students and, after reading it, he/she writes a note commenting on the students' arguments at the bottom of each composition. This can make students feel that their work has been valued and that someone paid it the necessary attention. Students, then, might feel motivated and dedicate more time and effort to their studies. Writing opinion notes to assignments is considered feedback and will probably influence students' willingness for future tasks.

In research on second language acquisition, the term feedback generally refers to the response that the listener or reader gives in relation to what the learner speaks or writes in the target language. This response can be a correction – corrective feedback –, or it can be a compliment, also called positive feedback (DULAY; BURT; KRASHEN, 1982).

3.3. Feedback and behaviorism

The concept of feedback was strongly present in the behaviorist model in the 1960s as a mechanism to reinforce desired behavior or discourage undesired behavior (MASON; BRUNING, 2001). This theory, based on behavioral psychology and structural linguistics, argues that learning a language is a completely external event, based on observable phenomena – the behaviors of the individual – what the organism says or does (PAIVA, 2014). Its main representatives were the psychologists Watson (1930) and Skinner (1992) as well as structuralist linguist Bloomfield (1933).

According to Skinner (1992), individuals only learned a language if they repeated their linguistic structures. The basis of this paradigm was the notion of stimulus, response, and reinforcement. The learners were exposed to external stimuli to which they produced a response and, if correct, this response received positive reinforcement. In the case of an undesired response, there was a weakening of that behavior until it no longer occurred.

For Watson (1930, p. 6), the language was nothing more than a simple type of behavior, which could be manipulated by means of conditioning. According to the author, a stimulus would be "any object in the general environment or any change in the tissues themselves due to the physiological condition of the animal". Likewise, his definition of response would be "anything the animal does." For this model of behavior to occur, the learner was always conditioned to respond correctly, because the exercise was designed in order to avoid errors and thus positive reinforcement was applied in sequence. In order to illustrate this idea, Watson (1930, p.6) argues "if you decide that the human organism should behave in this way, you must arrange situations of such and such kinds." In both a desired and undesired response, we can call this response as feedback of the individual. According to the behaviorist theory, reinforcement – or positive feedback – was what actually triggered learning (ANNET, 1969).

Although simplistic, such an explanation of this view of human verbal behavior shows that the learning of a language constituted a continuous process of external stimuli and responses meant to shape the structure of the language within the human mind. There are no references to mental processes, internal to the individual since the mind was seen as a black box, being impossible to deliberate on the events inside it.

The behavioral model is objective and rejects consciousness and subjectivism, considered abstract and irrelevant. Moreover, the error in the language being learned was an event to be immediately repressed and weakened not to compromise the formation of habits of the individual and the construction of the new language would be free from flaws.

3.4. Feedback and other paradigms

With the evolution of theories and models for language learning, especially with the emergence of Chomsky's generative paradigm in the late 1960s, along with the emergence of cognitive and constructivist theories, language learning has been interpreted as a mental process. One of the main ideas developed by Chomsky (1975) is that the individual detains a language acquisition device (hereinafter LAD), which is activated when the child begins to have contact with the language in childhood. This module would contain a universal grammar (henceforth UG) that upon activation would lead to language development. Under this proposal, grammatical structures are innate

to people, who are already born with the faculty of language, little depending on external stimuli of the environment in which they circulate. It is worth mentioning that the generative theory was primarily intended to explain the acquisition of the mother language (henceforth ML). However, many of its constructs have served as a basis for trying to understand this phenomenon from other angles, such as shifting the focus from the linguistic input to the linguistic creativity that learners have. This paradigm is concerned with unraveling the internal mechanisms of the mind instead of attributing the process of acquisition solely to external stimuli.

Faced with these transformations regarding the understanding of how to learn a language, errors and feedback had their status remodeled. The error started to be considered as an indication of the stages of development of the learner (ELLIS; BARKHUIZEN, 2005). Thus, there was room for the study of this phenomenon in depth in order to question, analyze and improve teaching methods. There was a substantial change in the perspective regarding learning. Researchers now observed a priori how and in which pace the individual acquired the forms of the language. This meant developing more effective instructional paths for learning to occur. Feedback, therefore, came to be seen as a relevant tool in helping language development, informing learners about necessary improvements in their interlanguage (MASON; BRUNING, 2001). Reinforcing this perception, Corder (1967 apud ELLIS; BARKHUZEIN, 2005, p. 51) points out that mistakes serve three purposes in language learning:

> [...]1) they serve a pedagogical purpose by showing teachers what students have learned that they have not yet mastered; 2) they serve a research purpose offering evidence of how languages are learned; and 3) they serve a learning purpose serving as an instrument so that students can discover the rules of the target language (eg, getting feedback on their mistakes).

Thus, feedback has taken a more significant role in language learning studies as it enables the improvement of the language being learned.

According to Paiva (2003, p. 221), "feedback plays an important role in human relationships, whether in spontaneous conversation, in classroom interaction, or in online interaction."[2] Feedback somehow generates interaction, since it guides the participants to walk a common path in communication so that there is an understanding

[2] Our translation. In the original paper: *"O feedback exerce um papel importante nas relações humanas, seja na conversa espontânea, seja na interação em sala de aula, ou na interação on-line."*

among the people involved. The lack of feedback can generate misunderstandings, gaps in the message flow, breakdowns in the communicative act, making it ineffective. We communicate in order to be understood and feedback regulates this process so that communication fulfills its purpose.

3.5. Feedback in the digital medium

There are many theories accounting for the process of learning a foreign language (henceforth FL). Among the most noteworthy ones are Watson and Skinner's behavioral model, Krashen's hypothesis, Long's interactionism (LIGHTBOWN, SPADA, 1993), cognitive theory (CHOMKSY, 1975), and Vygotsky's sociocultural theory (1991). It is worth mentioning these learning theories were developed at a time when the technology we have available today was not present in people's daily lives. Along with the innovations that have arisen, thinking, learning and teaching have undergone significant transformations, intertwined with these new tools that permeate society in the most diverse spheres of their lives.

Faced with this new reality, we understand that these constructs explain fragmented aspects about learning a language, and only one of these perspectives does not account for the process as a whole, which is complex and highly variable from individual to individual. There is still a need to seek a broader view, which considers both what is the innate and what comes from experience, to understand the learning process of a FL (LARSEN-FREEMAN, 1991). According to Paiva (2005, p. 23) and other authors (MCLAUGHLIN, 1987; ELLIS, 1990; BROWN 1993a, 1993b), the current models "do not include all the processes involved in the acquisition of a FL; I see them as fragmented visions from parts of the same system ".[3]

In addition to the fragmented vision for language acquisition and the urgency of improving pedagogical models for the digital age, the aspect of feedback in the development of a FL has always been an intricate matter. Much is invested in understanding brain functions and the amount of input to which the learner is exposed, as well as the influence of the relationships between individual and environment. However, feedback is relegated to a few pages in numerous pieces of research,

[3] Our translation. In the original paper: *"Os modelos de aquisição não contemplam todos os processos envolvidos na aquisição de uma língua e, muito menos, os de uma língua estrangeira (LE). Vejo esses modelos como visões fragmentadas de partes de um mesmo sistema."*

although it is strongly implicated in the interactions among learners – an essential component of successful learning (NAIDU, 2003). We can thus perceive that it is a moderately explored issue in view of other elements involved in language acquisition, which are extensively described and discussed. From this angle, we can understand the variety of existing, and sometimes divergent, denominations and categories in the literature concerning feedback. In the perception that what we had until then were viewpoints that needed to be understood as parts of a whole, the arrival of the digital era in a massive way was perhaps the missing element to transform this perspective. Therefore, we need to understand new conceptions that emerge concerning learning mediated by new artifacts and the implications of these on feedback.

Over the decades, the web has become an important means in the process of learning a FL. Through digital information and communication technologies (DICTs), the learners' exposure to the desired language has increased considerably, since there is no need for them to be in a classroom for this. Web 2.0, with its innumerable possibilities of making the user not only a mere consumer of information, but also an author of digital material (CARNEIRO et al, 2005), contributed significantly to making learning a language more practical and attractive. With video-sharing websites, textual production tools such as blogs and online dictionaries, and more recently educational software such as Duolingo, Babbel, Memrise, Busuu, among others, the users have contact with the language anywhere and anytime. From this perspective, DICTs provide a space for the development of autonomy and learners become fully responsible for their learning. Therefore, another view of this process is needed since both the way the linguistic content is presented to the user and the learner profile that consumes these digital products has transformed and language teaching and learning theories must somehow try to better account for this typical phenomenon of the 21st century. To corroborate this idea, Lévy (2004, p.7) states that

> [...]new ways of thinking and living are being developed in the world of telecommunications and information technology. The relations between men, work, and intelligence themselves depend, in fact, on the incessant metamorphosis of informational devices of all kinds. Writing, reading, vision, hearing, creation, learning are captured by ever more advanced computing.

Another issue should be observed more carefully when speaking about language learning through DICTs – automatic feedback. Much is discussed in terms of autonomous learners and in control of their development. However, as far as language

teaching applications are concerned, it is important to investigate how they offer feedback to their users who wish to enhance their knowledge in a FL as it directly influences the acquisition process.

Although new conceptions of feedback have arisen in the context of language teaching, Shute (2007) states that the feedback present in the tools still rests on a behaviorist view of teaching. Adapted to the digital age, feedback is often a message in audio, video, or writing in response to students' action to modify their behavior, shaping their perceptions. Thus, feedback in the automated digital environment can still be interpreted as a mechanism for behavior adjustment, as it was decades ago, failing to explore innumerable possible strategies for providing rich learning experiences.

There are several categories to account for feedback in the process of learning a language and in the construction of an autonomous learner. The literature of the area does not define a standard on these categories, mainly because feedback can be coined in relation to its content, its direction, and the moment it is provided or to its source (CARDOSO, 2011). Despite the typological heterogeneity, it is relevant to discuss some of these categories to understand in depth their objectives and how they operate.

According to Kielty (2004), there is recognition feedback, in which students receive feedback that their task has been completed. We can mention as an example the answer to an email saying "Ok, received." In this case, the interlocutor signals that the message has been received, but he does not comment on its content since he or she will probably approach it later.

There is motivational/interactive feedback, which is meant to provide the student with motivation and support to continue carrying out the activities (PAIVA, 2003). This type of feedback often appears in discussion forums, such as groups on the social network Facebook, created to complement face-to-face classroom, where participants need to interact with each other, reflect on a particular subject, and build knowledge in a collaborative way. In order for the debate not to be very brief, with little or no participation, the group mediator, usually a tutor or teacher, responds to each post by commenting on the subject and showing a positive reaction towards students' participation in the discussion, in order to keep them active and avoid the dissolution of the interaction. The following is an example of this type of feedback in a discussion

group used as a complement to discussions of the Educational Materials Authorship in 2016 course ("S" is the students and "T" is the teacher):

> S1: Hi, classmates
> I share with you the activity of memory.
> T: I liked the idea of listing the words students are going to use.
> S2: The activities are very cool, I'm sure they will contribute to the training and facilitate your life as coordinator. Is the last one (teaching grammar) ready?
> S3: Memory is one of the modules that I like best! Hehe :D[4]

It should be noted that this excerpt shows the presence of teacher comment soon after the initial posting, in order to show interest in what is being said, bring some consideration to the content posted by the student and thus motivate the other students to also interact.

Finally, there is evaluative/informative feedback (MASON; BRUNING, 2001; PAIVA, 2003), which is primarily related to showing the student some specific information and evaluating their answers. This type of feedback assesses students' tasks and offer complementary information, which can help them achieve the desired goal in relation to the task proposed. It is important to remember that this feedback can be provided not only by the teacher but also by a colleague or by the students themselves. When reflecting on and evaluating their own performance and needs, students can search for tools to help them in the process. Feedback can also be directed at an individual or a group, a teacher or a course as a whole.

Focusing on the level of complexity of the feedback offered, there is complex and non-complex feedback (SHUTE, 2007). The author explains that complex feedback is one that contains long and sometimes complicated messages that can cause distraction and confusion to the students. Non-complex feedback, on the other hand, has shorter, direct, and easy-to-understand messages.

Shute (2007) also addresses the terms synchronous and asynchronous feedback, which are related to the moment the user receives it: whether it is immediately or in a couple of minutes, hours or days after performing the online task. In chat rooms, we also perceive this synchrony, or simultaneity, as individuals are all communicating at the same time through the same tool. However, in forums or mailing

[4] Extract from a discussion in a private group on Facebook created by the professor of the course the researcher participated and its reproduction has been authorized.

lists, the feedback is seen as asynchronous, since there is a delay in the response to the executed action.

Five types of digital feedback were briefly discussed, focusing on different aspects involved in the teaching and learning languages scenario. As we can see, these categories are established according to three aspects: a) the existence of at least two interlocutors, b) the mediation of a machine, in this case, the computer, between the interlocutors, and c) the response time between the participants – synchronous or asynchronous.

3.6. Automated feedback

The types of feedback previously described demand the existence of at least two individuals in the digital teaching-learning process. However, in educational software for language teaching, we can observe that the provision of feedback does not require an individual on the other side of the device. This is what we call automatic or automated feedback, where the program itself receives and responds to information provided by the user.

Pyke and Sherlock (2010) call it technological feedback since it also has the potential to assist users with technical difficulties related to the platform they are using. It is very common in operational systems, such as Windows, when users are having problems connecting to the internet, for example, the system detects this malfunction, investigates possible causes and offers users some solutions. However, it is important to note that the technological feedback mentioned by the authors is not restricted to feedback produced by a program. The computer might only mediate it, that is, it can work as online support, in which users contact someone who can assist them in the technical problem via chat or email.

Filatro (2008) considers the term feedback as a broad concept that runs through various educational practices, from the control of the activities performed by students to the most personalized feedback. For her, the term ranges from instant, automated, standardized feedback produced by a program to a long-term one-to-one follow-up provided by a teacher or tutor on the other side of the screen. In any case, its traditional meaning in the field of Education is to return information to the student both during and after the completion of an online task.

Regarding automatic or automated feedback, the author considers it as the feedback that the computer gives to student responses during or after an online activity. According to her, this type of feedback, which is not the only possible digital medium, serves as user support to guide them in their learning process more directly. The author still offers options of this feedback model, such as feedback with right or wrong, messages of motivation or indication to complementary readings, which can improve deficits in the content addressed. It is valid to point out that these variations apply only in cases of closed answers, that is, when there is only one correct answer, as in objective tests (FILATRO, 2008).

Filatro (2008, p.130) established 8 possible levels of automatic feedback, depending on the type of learning involved that follows:

> 1. Indicate whether the answer is right or wrong, without any extra information.
> 2. Indicate whether the answer is right or wrong and explain why.
> 3. Provide resources so that students themselves determine if the answer is right or wrong and why.
> 4. Point out more appropriate strategies for the solution of a question, without specifying whether students are right or wrong.
> 5. Show students the consequences of their answers, especially with the use of games and simulations, in which each action is followed by feedback from the system.
> 6. Provide cumulative information about student progress during an activity - for example, reporting repeated error patterns or how close students are to achieving a pre-established criterion.
> 7. Record in photo or video demonstrations of psychomotor or affective learning, which should be observed by students individually or in a group, in order to verify step-by-step the effects of each action.
> 8. Provide extra activities so that students can apply the received feedback to new situations.[5]

[5] Our translation. In the original text: “1. Indicar se a resposta está certa ou errada, sem nenhuma informação extra.
2. Indicar se a resposta está certa ou errada e explicar por quê.
3. Fornecer subsídios para que o próprio aluno determine se a resposta está certa ou errada e por quê.
4. Apontar estratégias mais apropriadas para o encaminhamento de uma questão, sem explicitar se o aluno está certo ou errado.
5. Mostrar ao aluno as consequências de suas respostas, especialmente com o uso de jogos e simulações, nos quais cada ação é seguida por uma reação (feedback) do sistema.
6. Oferecer informação cumulativa sobre o progresso do aluno durante uma atividade – por exemplo, informar sobre padrões de erros repetidos ou quão próximo o aluno está de alcançar um critério preestabelecido.
7. Registrar em formas de foto ou vídeo demonstrações de aprendizagem psicomotora ou afetiva, que devem ser observadas pelo aluno individualmente ou em grupo, a fim de verificar passo-a-passo os efeitos de cada ação.
8. Oferecer atividades extras para que o aluno possa aplicar o feedback recebido a novas situações

In an attempt to illustrate the levels that the author mentions that are more identifiable in the context of language learning, we thought of examples so that one could better imagine the application of some of them in real contexts. For the first level, one can think of a simple questionnaire with multiple choice questions, where only one of the options is correct:

Question:

Check the correct option that completes the sentence below.

She's the beautiful girl.

( ) Are ( ) Is ( ) Am

When selecting one of the options, the tool, whether an application or website, signals learners response in green for correct response and red for incorrect, without presenting any additional information on the screen. The second level of feedback would bring an explanation accompanying students' response.

In case the answer is correct, a window might appear on the screen and provide the following message: "Correct answer. For he/she/it subjects the form of the adequate verb is 'is'". If students answered "are" the explanation could be: "'Are' is used only for the subjects you/we/they.". In the third level of feedback, the tool could present the student with a grammar table with the conjugation of the verb to be in the present tense by means of a button next to the exercise called "tip" or "help". In this way, the student would arrive at his own conclusions regarding the most appropriate answer. The table could be configured as follows:

I	am	a student.
He She It	is	an intelligent man. a busy doctor. an old dog.
We You They	are	good friends. very smart. strange people.

Table 2 - Example of the third level of feedback

Moving on to the fourth level, we also imagine the tool providing hints in the form of pop-up windows based on their incorrect answer, as in the sentence "She ___ a beautiful girl." Learners mark the answer "are" and when showing the error, offers the hint "We use a different form of the verb to be for people he/she/it". In the fifth level, there are the most varied online games aimed at language learning, such as Game Zone and Scrabble, which point out the real-time consequences that players' actions can cause. The last feedback would not be characterized as feedback itself, resembling more as an expansion of the practice of activities through hyperlinks offered to students after performing tasks and receiving feedback, as described by the author, so that they can use what they have learned in new exercises and reinforce knowledge. In her work, Filatro does not go in detail on this dynamics with extra activities. So, this is a possible interpretation of the eighth level of feedback she brings.

For more complex activities involving mastery of higher-level cognitive abilities, and that can hold several solutions, the author suggests that feedback be provided during the exercise execution rather than just after it is completed. In addition, it is important that feedback is focused on processes in the same way as for the results of tasks. This way, students have guidance on their learning process about the mistakes they made and their causes, affording moments of self-assessment and reorientation of their studies.

However, it is important to point out the fact that activities with a variety of possible responses, that is, with open answers, the feedback becomes unpredictable, and it is not possible to fully anticipate what learners will respond, and thus the automation of feedback becomes more difficult. In order to deal with the unpredictable, an educator must be available, the one who will be responsible for producing feedback according to the specific needs of the learner, filling the gap that programming is not yet able to do. Program algorithms of this kind have not yet reached this level of individualization and contextualization to the point of offering feedback to any possible response produced by the student. This technological limitation brings the need to understand more broadly the question of the interactivity of the educational programs for electronic learning, a subject which will be discussed on the following section.

While Filatro (2008) delimitates in detail possible levels of automatic feedback considering several aspects, it is important to observe that they are not specific to the scope of language learning. However, Leffa (2003) brings automatic feedback specifically directed to the teaching of FL. According to him, corroborating with the idea

of Filatro (2008), in order to feedback fulfill its function more efficiently, it must be diversified. This means that there must be different types of feedback according to the desired goal in the task. Thus, the author created three categories: a) generic, b) situated and c) strategic. The first is the simplest, only showing users if their answer is correct or incorrect. The second is a specific comment made on the response that users give, simulating a possible face-to-face situation. Feedback can be both corrective and positive. When correction is necessary, the comment is based on the students' response, attempting to induce students to seek the correct answer by means of a hint, of what is missing or what is incoherent. When feedback is positive, the comment consists of motivational words. The third type of feedback, however, focuses on the search for learning strategies, in which the answer is not given directly, but tries to show the way for learners to reach it. In the words of the author (LEFFA, 2003, p. 38):

> [...]while generic feedback is in the right or wrong assessment of the response, and the situated feedback shows the source of the problem, strategic feedback attempts to suggest learning strategies that can lead the student to the right answer. You do not give the answer to the student, but try to show him how to get to it.[6]

By observing the categories of automatic feedback described by Leffa (2003) and Filatro (2008) one can notice similarities despite different nomenclatures. The generic feedback proposed by Leffa is in line with the first level of feedback that Filatro lists in her work since both of them highlight the issue of presenting results only, with no additional information offered to learners. The situated feedback proposed by Leffa is close to the third level of feedback presented by Filatro since both of them mention not delivering the correct answer directly, but provide ways or tips for students to be able to find it for themselves. Finally, Leffa brings the strategic feedback, which resembles the fourth level of feedback described by Filatro, since both focus on providing learning strategies so that learners can develop reasoning and reach the expected response in the activity.

In line with Leffa (2003), Horton (2000, p. 27) also states that the simpler the feedback the less effective it becomes. Feedback such as "right" or "wrong" does not

---

[6] Our translation. In the original paper: "*Enquanto o feedback genérico fica na avaliação da resposta, certa ou errada, e o situado mostra a origem do problema, o feedback estratégico tenta sugerir estratégias de aprendizagem que possam levar o aluno à resposta certa. Não se dá a resposta ao aluno, mas tenta-se mostrar-lhe como chegar a ela*."

motivate students. The author also suggests that for a correct answer, the feedback should contain an explanation of why it is right and for an incorrect answer, feedback should be kind but at the same time clearly show the error and offer tips for users to reach the correct answer.

Vetromille-Castro (2003, p. 14) also makes suggestions regarding automatic feedback in language teaching when he says that feedback guides and motivates learners. He states that for this reason, feedback is essential because learners cannot count on the teacher in the same way as in face-to-face teaching.

> [...] feedback in a virtual environment has to show the student that the path (s) he/she is following is correct, clarify doubts about content and use of the system, and provide, when necessary, pedagogical guidance.[7]

Automatic feedback is still a rich field to be explored when it comes to technology-mediated language teaching and learning. This fact is evident when we enter the term on Google search engine, for example, and the results are, for the most part, other areas of study, such as Medicine and Informatics. When we search for databases with more academic support, the result is not much different. If on the one hand references in the area of Applied Linguistics are still incipient, on the other hand, this scenario makes the field of research fertile to new perspectives.

3.7. Formative and summative automatic feedback

One of the foundations for producing meaningful feedback in the e-learning literature is one that provides a means of working the needs of learners, providing input so that they can build knowledge and not only be given a grade on their performance. A relevant vision for feedback elaboration in order to provoke the expected results is that learning is a process rather than a product. Feedback must work as a trigger, leading to moments of reflection along the way and consequently fostering learning meaningfully.

To consider whether feedback goes beyond a right or wrong signaling and/or scoring activities, a practice still very common in electronic learning, it is opportune to

[7] Our translation. In the original paper: “[...] o feedback no ambiente virtual tem de mostrar para o aluno que o(s) caminho(s) que está seguindo é(são) correto(s), esclarecer dúvidas de conteúdo e de uso do sistema, além de fornecer, quando necessário, orientação pedagógica.

establish the concepts of formative and summative feedback to better understand the effectiveness of their use in virtual learning tools. It is paramount to understand how an automated learning environment works on improving writing in a FL using programmed responses built from an algorithm.

Formative feedback is meant to instruct learners on how to acquire the knowledge needed for a given task rather than inform them of the accuracy of their response. It is multidimensional, non-evaluative, supportive, specific, timely, believable, infrequent, and genuine (BROPHY, 1981; SCHWARTZ, WHITE, 2000). Feedback, then, becomes an instructional device, which should aim to teach something new and add unprecedented information to what learners already know. This view of feedback as an instructional mechanism is brought by Hattie and Timperley (2018) and relates closely to the constructivist view of learning, as discussed by Vygostky (1991). From this point of view, feedback provides new information that merges into learners' intellectual baggage, giving them support to reach their full potential, which results in an assisted performance. This way, learners build knowledge in a process of scaffolding, in which they explore their zone of proximal development (hereafter ZPD). Using this assistance or instruction, learners are guided in a way they achieve performance in a given task that probably they would not be able to reach by themselves at that particular moment of their linguistic capacities.

Shute (2007) has addressed in depth this category of feedback directed to learning in virtual environments and her work serves as a reference for this section. According to the author, "formative feedback represents information communicated to the learner that is intended to modify the learner's thinking or behavior for the purpose of improving learning." (SHUTE, 2007, p. 9). For this to happen, firstly, feedback needs to bring solutions without raising uncertainties. In this sense, feedback is considered more effective when it provides details on what can be improved in learners' response rather than when it shows only whether it is correct or not. The lack of feedback specificity can lead to frustration, compromising the motivation to use the virtual learning tool, since learners may feel that it is their inability to understand the feedback when, actually, the feedback is little elaborated for the objective of the activity (KLUGER; DENISI, 1996).

Given this premise, the ideal feedback needs to present two crucial elements: verification and elaboration (MASON; BRUNING, 2001). The verification, also known in the literature as knowledge of results (hereinafter KR), is meant to evaluate the

response as correct or incorrect. In other words, the system generates a reaction to the learners' response in audio or video or a combination of the two, which is presented on the screen without further details.

Elaboration, on the other hand, is concerned with providing useful resources to learners. Thus, there are messages that offer examples of what should be corrected, discussion about the mistake made, why it occurred, and gentle, welcoming tips that will lead learners to repair it. It seems clear that the functioning of formative feedback, in general, should be positive, generating feelings that inspire motivation, self-esteem, safety, and credit in learners as opposed to feelings related to testing, elimination, and intellectual incapacity. It is important to note that with this type of feedback there is no concern in providing correct answers. Learners themselves must find them and this becomes a safe process when feedback is offered with facilitating elements. In line with this view, Alessi and Trollip (2001, p. 115) state that

> The importance of feedback to all computer-based learning environments is widely acknowledged: 'Feedback should be positive. It should avoid negative statements, sarcasm, and should never demean the learner. Feedback should be corrective. It should provide the learner with information to improve future performance.

Seeing that formative feedback configuration is tangent to the constructivist teaching-learning perspective, it aims to assist learners until they are able to reach conclusions on the subject studied by themselves. When you are in an environment in which the interaction takes place among people, it seems that the process unfolds without great obstacles. However, in the virtual environment, in a model of interaction between learners and software, there are some setbacks to build an efficient process. Broadly, these constraints happen mainly due to programming limitations that compromise the processing of natural language at a level that meets the needs of the learners.

Adding personality and learning style to mouse clicks and window cascade is not yet a palpable reality. In this way, a program focused on teaching and learning should display elements that guide users' actions and then be interpreted by the system in order to offer timely support. Then, uniformity hinders possibilities that would perhaps be infinite in human-human dynamics. That being said, the program design is what establishes how the interaction and feedback dynamics happen. Its agency extends only to the measure of human need, that is, its power of creativity regarding the input of learners is finite. Without prior programming of the results desired,

computer-mediated learning becomes brief and superficial and therefore is little helpful in pedagogical terms.

The work to develop tools that more and more meet human needs in an autonomous learning context is intensified when there is a reference that computerized feedback interventions, which are predominantly in visual/written format, have more positive effects than non-computerized interventions (KLUGER; DENISI, 1996). The sense of impartiality generated by interaction with an inanimate agent engenders learning situations distinct from those found in face-to-face events involving educators and learners. In this scenario, another favorable factor is the instantaneity of delivery, especially in the age of digital learners, with increasingly brief attention span. In addition, the possibility of a higher frequency of feedback due to its configuration and automation also corroborates to a positive effect of computerized responses. According to Hurd (2007), the lack of instant feedback is the most common cause of anxiety in an autonomous learning environment, since learners experience a sense of insecurity as they are making decisions without the frequent follow-up of someone more experienced.

The general guidelines that Shute (2007) treat as primordial for formative feedback to have effectiveness are part of a process and not a static, closed construct in itself. The author divides the guidelines into four aspects: a) things to be done; b) things to avoid; c) timing issues and d) learner characteristics. Regarding automated feedback, the time issue does not apply, since all feedback is provided immediately. Regarding the learners characteristics, automatic writing evaluation programs still do not have the technology to perform categorization of learners' profile in order to direct the feedback, so this aspect is also not applicable to the scope outlined in this investigation. The item "b" brings a wide range of feedback delivery, such as oral and audiovisual, which are not present in the program under scrutiny. Thus, the core evaluation of feedback focus of this research remains in the items prescribed in the first aspect, and are listed below in nine topics:

Things to do:

1. Focus feedback on the task rather not on the learner;
2. Provide elaborate feedback to enhance learning;
3. Present elaborated feedback in manageable units;
4. Be specific and clear with feedback messages;
5. Keep feedback as simple as possible but no simpler (based on learner needs and instructional constraints);
6. Reduce uncertainty between performance and goals;
7. Give unbiased, objective feedback, written or via computer;

8. Promote a learning goal orientation via feedback,
9. Provide feedback after learners have attempted a solution.

It is challenging to gather solid foundations in a coherent way for a theme still in abundant development. Coupled with the fact that online e-learning programs are launched or updated with considerable frequency, it becomes an intricate process to provide a theoretical basis for fail-proof and sufficiently comprehensive to innovations in the field. A priori, the main objective is to examine the program mainly in its human-machine dynamics and how this relationship influences the construction, in the Vygotskyan term, of the writing skill in English to the extent that this scenario is viewed from a post-human perspective, in which objects become agents more than just mediators in this learning configuration.

According to Hattie et al (2007, p. 8) "feedback is configured as information coming from an agent, be it a teacher, a colleague, a book, a relative, himself or experience." If a computer program is able to perform this action, then, it is also considered an agent itself. That being said, it is imperative to design appropriate constructs to embrace this new perspective in which the thought does not concentrate solely within the human mind any more, but it is allocated in different artifacts that compose the context in which one lives and, thus, reshape core aspects concerning language acquisition (PENNYCOOK, 2018).

Still, we must consider that some of the properties presented and combined in this section may not completely match the examination that this work intends since there is not yet a specific theoretical framework to approach this type of system. We assume, then, the risk of not being able to encompass and analyze some properties in their entirety. This research, due to the aforementioned conjuncture, may contribute to a potential set of directives conducive to the development of future online tools for FL writing skills.

3.8. Feedback and written communicative strategies

Based on a constructivist, complex and posthuman view that permeates this work, written communication strategies must be triggered in the learning process when feedback is efficient. Since automatic feedback is a field still under development in the area of Applied Linguistics, we need to pay attention to what strategies are incited in

this computer-mediated autonomous learning fostered by a human-machine relation and to observe the way learner support is characterized.

The literature of the area mainly includes studies on communication strategies (hereinafter CS) focused on learners' oral production, which are based on face-to-face interactions (TARONE, 1977; KUMARAVADIVELU, 1988; DÖRNYEI, 1995). The term CS was coined by Selinker (1972) in his seminal work on interlanguage, in which he considers them one of the five mental processes involved in the acquisition of FL. CSs are a reflection of learners' conscious attempt to solve problems in communication when their linguistic ability is insufficient to convey the intended message.

Canale and Swain (1980, p. 30) also bring this matter into their model of communicative competence by naming it strategic competence. The authors define it as "verbal and non-verbal communication strategies triggered to compensate for communication failures due to variables involved in the performance or insufficient competence." Using the term "non-verbal" in their definition, we understand the focus is on oral language rather than writing. In general, the several definitions that have emerged over time in studies on the subject take speech as the expression of the language under study.

There are different terminologies available to classify and explain CS. In general, they treat the message as a whole rather than isolated words. However, in written language, problems can be allocated in isolated terms, such as incorrect spelling and word order, errors of agreement and verbal inflection, etc. Thus, it is understood that without specific studies of CS when used to deal with errors in the cohesion of a text raised by automated feedback, adaptations are necessary to better account for the data. This limitation is brought in the work of Kellerman et al (1987) in the sense that the various terminologies coined for strategies fail to address their use in the written mode.

Communication strategies basically contemplate two perspectives: a) change, reduce or avoid an idea due to lack of conditions to express it, and b) manipulate the available linguistic knowledge to try to convey the message despite the obstacles (FÆRCH; KASPER, 1983). In order to relate the existing taxonomies, Dornyei and Scott (1995) combined the different current classifications. The authors outlined an extensive list covering 33 strategies. Of these, only 13 are suitable for use in writing, which are: abandonment, reduction, omission, substitution, all-purpose words, (perspective "a"); and paraphrase, approximation, restructuring, literal translation,

foreignizing, word coinage, code switch (perspective "b"). Although this classification is based on learners' oral production, we believe these CS can account for occurrences in both oral and written form.

Moattarian and Tahririan (2013) investigated Iranian EFL learners' use of CSs in oral and written performances at two levels of proficiency. To this end, 60 university students of EFL were selected and assigned to two distinct groups. The participants' oral and written performances were analyzed quantitatively and qualitatively using Dornyei's (1995) taxonomy of CSs. Their descriptive study showed that learners employed CSs to compensate for three main gaps: lexical deficiency, problems in discourse management, and uncertainty in conveying the message. To compensate for lexical deficiencies, the participants used approximation, code switching, circumlocution, word coinage, use of all-purpose words, and foreignizing. For deficiencies in discourse management, participants adopted paraphrasing, and message abandonment. The third gap regarding uncertainty in conveying the message was observed only in oral performance, therefore, there were no results oriented to this gap in writing.

From this same study, another relevant conclusion was drawn regarding the amount of CSs employed either in oral or written language. Learners used more CSs in their oral production rather than in their written production, fact that calls the attention to the necessity of previous instruction on the strategies learners can make use to develop their writing skill consistently. The authors also stated that the gaps can be considered by language teachers and material designers in order to address learners' limitations more efficiently.

A study focused on the use of CSs in 40 Turkish students studying English Language at university level found that the omission was the most used CS in written assignments (ELYILDIRIM, 2017). The investigation was based on the comparison of writing in the ML and FL about the same topic, which evidenced the omission of details in the text produced in the FL, which were previously written in the ML. The decision for applying such strategy points to a tendency among learners to avoid taking risks since they do not try to expose their ideas when they do not have enough linguistic competence. Another finding on the use of omission is the lack of interest in venturing into other ways of expressing themselves, possibly because it requires more effort to try to convey the original message in other words.

At the level of interaction, which is established in the communication between individuals with a common purpose, the use of omission can be due simply to lack of need to express ideas clearly and completely in a non-authentic context. In our 15 years of experience in English teaching, the practice of writing skills in a formal learning environment, in general, is based on the execution of tasks by learners regarding topics chosen by the teacher/tutor/instructor to keep track of their development in a certain study program. In this perspective, writing moves away from its dialogical character, making its implementation not a stimulating enterprise. The commitment to be trustworthy with words is sometimes not fueled by a motivation that goes beyond accomplishing a task to fulfill curricular obligations. The inauthenticity of the writing practice context discourages learners' responsiveness since the interlocutor is often fictitious. Therefore, a more convenient strategy becomes the one that requires less effort to be complete, such as omission.

Restructuring is another usual written communication strategy used when learners rely on the possibility of manipulating the intended message (the one they have primarily in their minds or in a text) using words and structures they feel confident about. Larios et al (1999, p. 16) defines restructuring as "the search for an alternative syntactic plan once the writer predicts, anticipates, or realizes the original plan is not going to be satisfactory for a variety of linguistic, ideational or textual reasons." The ideia of restructuring can be understood beyond written compositions in a FL. The author of the study explains that if we take into account Piagetian view of cognitive development, going from one stage to another, restructuring is in the basis of human cognition evolution. By the time new input comes in, the cognitive system needs to restructure itself in order to accommodate it and build new connections. Regarding language, the individual must channel his/her ideas through the interlanguage he or she has available at the time. Restructuring, then, becomes essential since the thought must be accommodated in the structures the individual can manipulate, which sometimes is unsatisfactory. An example of the evidence of restructuring is provided in figure 3:

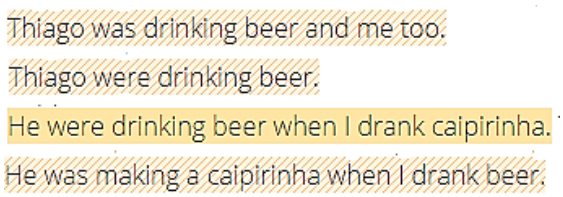


Figure 7 - Example of restructuring strategy

The author conducted a study with five Spanish speakers learning English in higher education as a foreign language and found that learners use restructuring for two different objectives. Besides compensating for lack of repertoire in the TL, this strategy can be applied as a means of expressing a writing style. The two uses were directly related to the students' proficiency level. The intermediate level learners used restructuring mainly for compensatory purposes, such as lack of lexical or morphosyntactic items, while the advanced level learners used it to avoid lexical repetition or to have a better match between intended and expressed message.

Since the aforementioned study was conducted using pen and paper and not in an online tutoring system with automatic feedback, the idea of using restructuring for stylistic reasons might not apply for the work being described here. One reason is related to students' proficiency level who used this written communication strategy. They were advanced students, which means they probably had a larger linguistic repertoire at hand to manipulate the texts to suit stylistic purposes. The subjects in this present investigation were beginner and intermediate levels. The other reason is the composition process being mediated by the feedback provided by the tutoring system, which provides correction and guidance merely at the surface level, ignoring any creativity or style of the writer. In view of the above, the restructuring strategy might be evidenced differently because of the conditions in which it occurs.

Chimbganda (2000) investigated the use of CSs by university students of Biology. He found that students were eager to use L2-based strategies as circumlocution, and paraphrase and concluded that those who took the risk of applying resource expansion strategies irrespective of grammatical problems were more successful in achieving their goal of communication. The same way, Aliakbari and Karimi (2009) investigated the use of CSs in the written performances of EFL learners at different proficiency levels. They found that the higher the proficiency level, the more reconceptualization strategies and the less substitution strategies were used. That means that insisting on transmitting the intended message yielded better results in writing than avoiding it because of possible linguistic limitations.

In the area of CALL, Chen et al (2015) developed a corpus-based automated assistant to help Chinese speakers learning EFL with the use of paraphrasing, another communication strategy applied to writing when one faces linguistic limitations.

Although the use of a thesaurus seems to offer the same assistance, PREFER (the acronym for PREFabricated Expression Recognizer) focuses on phrasal paraphrases rather than only lexical paraphrases. 64 Chinese EFL speaking college freshmen in a public university of an Asian country were recruited for participation in a study to test the efficacy of PREFER. The results showed that, compared to a thesaurus or an online dictionary, the paraphrasing tool managed to better satisfy learners' needs, specially in phrases. Also, the less proficient and more motivated students were able to produce good quality paraphrases and in great quantity with the help of PREFER. The results shown with this tool can provide insights on the need for corpus-based automated systems that gather structures from different varieties of English. English is not exactly the same everywhere and this characteristic must be taken into account when developing applications to assist learning a lingua franca as this language has become throughout the centuries (CANAGARAJAH, 2007).

As the interest of this research is the investigation of how an automated assistive tool contributes to the improvement of EFL learners' writing skill, we must limit the scope of CSs in relation to those possible in these conditions. We understand that the use of strategies to solve problems detected in a text by automatic feedback evidences its effectiveness on assisting learners to better communicate their message. The learners' reaction to feedback can be understood by identifying the strategy that they use to solve the problem. If the strategy used addresses the problem satisfactorily, we understand that feedback has served its purpose, contributing to the development of learners' skill in a positive way.

4 Interaction and interactivity

4.1. Interaction

In light of language learning in digital environments, there are two relevant aspects to consider when thinking about the success of the process in this scenario – interaction and interactivity. It is believed that for the learning of a language to happen effectively, individual and virtual environment need to come into contact, and the contact needs to be defined in the sense of who or what is interacting and at what levels and conditions this happens. In order to better understand these concepts, it is important to first establish how they have been recognized within the most widespread learning theories over the decades.

By the time behaviorism was in vogue, it was understood that the interaction of the individual with the environment was what made learning happen by shaping that environment so that it better provides conditions (stimuli) for the individual to produce the desired response. This learning model is considered objectivist, in the sense that it sees individual and environment, mind and reality as separate entities and independent from each other. In the context of learning, one has the vision that knowledge is external and objective and there are linear paths, organized in stages to reach it (QUADROS; FINGER, 2008). The interactions, from this perspective, are meant to take learners mechanically along the correct path so that they are able, in the end, to demonstrate that learning has been consolidated. Principles such as negotiation, discovery, and sharing of information are not congruent with this learning model and, therefore, it is beneficial in situations that do not require decision-making and problem solving, which require an evaluative analysis of data, for example.

In the '60s, we have the cognitive view, where reason is what leads to learning, so it is internal to the individual and the environment little interferes in the mental processes of knowledge acquisition. This pragmatic approach to learning is supported by Piagetian studies on the development of the human intellect. According to Ellis (1999, p.17, emphasis in the original) "the role of interaction is to supply the "black box" of the mind with the correct type of data for internal mechanisms to manipulate them". In this way, new elements interact with the prior knowledge of the individual within their mind, generating new knowledge, what Piaget (1973) calls the process of equilibration. This individualized and initially biological event is intrapersonal, since it occurs without any kind of intermediation with the other extreme, whether it is another

person or another object present in the world. It is worth mentioning that under the piagetian view, social interaction is not ignored. However, it has a supplementary status, increasing and enriching the learning structures (FERREIRA, 2003).

Years later, in light of Vygotsky's constructivism (1991), when learning, the individuals filter the experiences they have throughout life to construct their own interpretation of reality. There is the creation of meaning by opposing the idea of its mere acquisition by individuals. That said, interaction has a social bias, and this dynamics is only possible among individuals sharing time, manipulating information, and building knowledge together. An object, such as a smartphone or a computer, would never constitute the other extreme of the interaction, being only a mediating instrument for the interaction act to materialize itself. According to Lantolf (2000), one of the key concepts in the constructivist view is the idea of mediation, in which the individual only reaches more complex levels of thought when interacting with other individuals through the language they have in common. This interaction has the purpose of assisting the less able one to develop abilities with the aid of someone more capable at the moment – the scaffolding – also a key principle in this interpretative perspective of learning.

The humanistic view in the scope of language learning manifested itself through the work of Maslow (1971) and Rogers (1980) and brought a different interpretation in relation to the purpose of the interaction among individuals. According to the authors, teachers and students interact on an equal level and the teacher acts as a facilitator, respecting their time and providing a friendly environment for learning to happen. In this sense, the focus was on the affective aspect of learners, which is composed of emotions that must be taken into account in a teaching-learning scenario. Other constructs relevant to the humanist perspective are the issue of empowerment, in which learning must lead individuals to self-knowledge in order to reach their maximum level of development regardless of inherited, historical or environmental limitations (SHIRKHANI; ARDESHIR, 2013). Human interaction is a basic requirement for the development of individuals as a whole, for they are social in their very essence.

At the time the humanist paradigm was in vogue, human-machine interaction was not a topic of discussion as it is today. In the 1980s, computer programs did not yet have a massive presence in learning environments, and comprehension on the impact of this new configuration of relationships mediated by portable and intelligent technologies was scarce. However, the 21st century has brought a reality in which

people and machines are in symbiosis and their interaction is a relevant theme to understand what elements cause effects in cognitive processes and understanding and acting in society. The time people spend doing on-screen activities has increased considerably, and the human replacement we observe should be considered alarming. Currently available technologies have been reducing opportunities for people's interactions with some degree of criticism (HIRSH-PASEK et al, 2018).

Taking this observation into the teaching-learning context, learning a language more autonomously using an automated program that offers few opportunities to develop critical sense can pose some disadvantages if the purpose of the technological apparatus is to go beyond the decoding of the language structure in question. Yet, even with the primary task of unraveling the rules and uses of the language being learned, mastering it without knowing what it entails in the world can restrict the development of the learners' symbolic competence, who become unable to understand and be part of the socio-historical-political-cultural power game that language is a constituent element (KRAMSCH, 2015).

The different perspectives for the concept of interaction are consonant with the philosophical-educational paradigms that prevail in each time. However, the digital age drew attention to a review of the basis by which the teaching and learning processes were thought until then since there is growing evidence that the processing of information by the human mind has undergone significant transformations with the advent of computerization and its technological artifacts (PRENSKY, 2001; BROWN, 2002; BARONE, 2003; SIEMENS, 2004). Thus, it is of fundamental importance to understand where and how the digital medium, which materializes itself in electronic devices, is positioned in this spectrum. Moreover, we must evaluate the impact of human-machine interaction in learning in times when the virtual abstraction takes spaces previously essentially taken by humans and objects not artificially intelligent, as well as how it is going reshape our view of ourselves and our doing in society.

Facing the belief in the need for a transition to accompany this new context of teaching and learning, Siemens (2004) proposed the connectivism, a learning theory that takes into account this new format of reality, in which people are constantly connected by virtual tools, which expand the reach to information at speeds formerly unseen and, therefore, provide different moments and spaces for the construction of knowledge. In the words of the author "[...] technology is altering (reconnecting) our brains. The tools we use shape our thinking." The goal in developing this theory would

be to counter traditional formal education found in institutions, which are far from this new dynamics favored by technology present in everyday life.

For connectivism, the social is created and recreated in a constant flux, non-linearly, through connections that move between individuals repeatedly and forming networks. According to Barabasi (2002, 2010), the networks, which are formed from these connections – interactions – are all over the place, we only need to be able to see them, implicit in today's society. In addition, the epistemological basis of this theoretical view led to the elaboration of Downes' (2005) proposal of distributed knowledge, in which knowledge is not contained in any physical and concrete place; it is distributed among the members of a connected group sharing common characteristics. The connections among individuals are, from this perspective, knowledge itself, and can not be detected in its totality in a single individual, but manifesting through interactions manifested among creatures.

However, this new theoretical view has elicited some criticism for not being able to substantially replace the knowledge already reached in terms of how learning happens. Kerr (2006) conveys the opposite idea when he states that technologies have not transformed what is meant by learning to the point where we can ignore the previously established paradigms on the subject to propose a new theory. According to the author, with a more observant look, the idea of non-linearity and connectivity was once addressed by the Chaos/Complexity theory applied to language learning, for example. From this frame of reference, everything changes at any moment because people are affected by the environment in a continuous flow of interactions, unpredictable and non-linear, generating transformations and adaptations at different levels, causing or not learning (BERTALANFFY, 1973; LEWIN, 1994; Larsen-Freeman, 1997; PAIVA, 2011).

Having said that, it is imperative to develop a theoretical basis that can embrace with more comprehensiveness and depth how people access information, build knowledge and learn in the digital age, especially with regard to the human-machine relationship, increasingly interactive and automated. In order to corroborate the thought that perhaps connectivism is not something totally original in dealing with learning in the digital age in a different way, Kopp and Hill (2008, p.11) state that

> A paradigm shift, indeed, may be occurring in educational theory, and a new epistemology may be emerging, but it does not seem that connectivism’s contributions to the new paradigm warrant it being treated as a separate learning theory in and of its own right. Connectivism, however, continues to

> play an important role in the development and emergence of new pedagogies, where control is shifting from the tutor to an increasingly more autonomous learner.

4.2. Interactivity

In the field of informatics and instructional design, the term interactivity is very widespread, but this does not spare us from having several conceptions and definitions, which sometimes brings about some disagreements regarding the establishment of a solid paradigm in the area. In general lines, interactivity is characterized by the contact between the individual and the object. In this conception, it is understood that the other extreme of the dynamics does not imply the presence of a subject (SMITH, 1983 apud SIMS, 1998). Thus, we can have as an example the dynamics between a person and a computer program. This program, embodied in electronic equipment is the other end of the dynamics and not just a mediating tool. In this scenario, what mediates this contact is the language they both share, because the program can manifest itself on the computer screen in written, imagetic and auditory form and thus the exchange can be established. This vision is in line with vygotskian language thinking, which sees language as a cultural artifact mediating psychological and social activities and can also be applied to the process of foreign language acquisition (PAIVA, 2009).

In the field of sociology, Jensenn (1998, p. 188) suggests the definition that interactivity is "[…] the relationship between two or more people who, in a given situation, mutually adapt their behavior and actions to each other." In the scope of the human-machine relationship or human-computer interface, the determinant of interactivity is the "[…] the style of control that exists between the human and the computer is the key determinant to interactivity.” The author emphasizes that the interactivity, depending on the focus that it imposes, can hold several definitions, not being a monolithic concept but rather multifaceted. In the same line, Wiener (1999, p. 257) states “control in other words, is nothing but the sending of messages which effectively change the behavior of the recipient”, definition which shows how significant it is for interactivity to promote the possibility of reacting to stimuli.

It is not uncommon the exchange of the terms interaction and interactivity in authors who write about distance learning and e-learning. However, the distinction is relevant in view of the direction to which this work is committed to seeking to

understand the manifestation of these constructs in an automated application for the learning of writing in FL. Thus, Filatro (2008, p.107) provides a definition for the terms in the scope of instructional design in which "the interaction concerns the behavior of people towards other people and systems. Interactivity, in turn, when describing the ability or potential of a system to provide interaction, is a prerequisite for interaction." Still in the light of interactive technologies, Tori (2010, p. 5) has the following definitions:

> Interaction: action exerted between two elements, in which there is mutual interference in the behavior of the interactors. Interactivity: perception of the capacity, or potential, of interaction provided by a given system or activity. In this way, "interaction" will be considered as the interactive action itself, while "interactivity" will be treated as a property of the environment, technology, system or activity.[8]

In a more unbiased context, according to the dictionary of Portuguese Aurélio Online, interactivity is a "faculty of exchange between the user of a computer system and the machine, through a terminal equipped with a screen" and an "interactive media character". According to the English-language dictionary Cambridge Online, interactivity is defined as "the involvement of users in the exchange of information with computers and the degree to which this happens." From these meanings, it is clear the need for the presence of a computer or something similar so that an interactive event is constituted as such. Also, it is noticed that the term exchange is present in both definitions, suggesting that there is indeed a movement between the elements involved, in which case information is shared, generating a favorable scenario to learning.

With regard to the question of the level at which involvement occurs, an idea present in the definition of the Cambridge dictionary, we can see the indication of measures or conditions that provide interactivity. We understand, then, that not only the user of a program developed for language teaching, for example, learns a language as well as the program itself recognizes patterns in everything that user executes in it and, consequently, assimilates information and builds new interpretations, systematizing new patterns in a continuous process (FILATRO, 2008). Interactivity allows users to instruct the system through their actions leading to its transformation and adaptation, which, in turn, creates new situations for users in a constant feedback loop.

8 Our translation. In the original: “Interação: ação exercida entre dois elementos, na qual haja interferência mútua no comportamento dos interatores. Interatividade: percepção da capacidade, ou potencial de interação propiciada por determinado sistema ou atividade.”

In the midst of several definitions of several areas of knowledge and the dynamicity that interactivity can provide in the human-machine relationship in a system, it is imperative to clarify the aspect of intentionality in the user-content interaction. Although there is room for one to influence and modify the other, interactivity is a planned feature in nature, that is, there is a need for an agent to visualize, consider a range of possibilities and program them into the system. Unpredictability is not yet compatible with automated learning tools. This is one of the limitations of computer programs once they need to be previously fed through a database so that later they can provide degrees of interactivity to users. This means that unscheduled and unplanned questions or actions not fed into the system beforehand remain unanswered or with no feedback. In the words of Leffa (2006, p. 194), "although the computer is capable of generating and managing its own unpredictability, starting from its own algorithms, it is unable to improvise and interact beyond the parameters introduced in its programming."[9] The author further complements the idea by stating that a good, yet automated, learning-oriented program at some point must enable the interaction between learner and teacher/tutor/instructor because only then it shifts from a mass communication medium to a system with some degree of customization that meets individual needs.

According to Rhodes and Azbell (1985), interactivity can be classified according to the relationship between human and machine, in the sense of how much autonomy the user and the program have. There are three interactivity degrees: a) reactive, where the user has restricted control over the virtual content, with options and feedback pre-established by the program; b) Coactive, in which the user has control over the sequence, rhythm, and style of the actions performed in the virtual environment and c) proactive, in which the user has control over the content and structure of the program.

It is important to point out that the reactive and coactive categories can be seen as an interpretation of the behaviorist learning theory, considering that it is the same mechanism used in the teaching machine created by the psychologist Sydney Pressey in 1925 (PITARELLO, 2014). Both technologies – the Pressey machine and a computer program – though almost a century apart, are often based on the concept of

[9] Our translation. In the original: “Embora o computador seja capaz de gerar e administrar sua própria imprevisibilidade, partindo de seus próprios algoritmos, ele é incapaz de improvisar e interagir além dos parâmetros introduzidos em sua programação.”

stimulus-response between the individual and the apparatus. This approximation occurs because the behaviorist theory goes in line with instructional design in the sense that the computational programming is also binary itself, with zeros and ones, if and else, on and off, and so on. The question that emerges from this similarity is that language learning is not limited to a dual set of answers, right and wrong, yes and no. Thus, we observe that a program designed to teach some aspect in the scope of languages following only this theoretical model becomes restricted in affording rich moments for learning when circumstances require more elaboration and reasoning.

The more proactive interactivity in an e-learning space offers, the more flexible it is, and the more likely it is to sort out unpredictable responses inherent to human behavior. The less rigid it is, the more the material approaches the constructivist paradigm, where the path is structured along the process, offering users access to different cognitive stimuli and assisting them in the elaboration of new knowledge through less programmed sequences. Dickinson (1995, p. 145) reminds us how interactivity can be more than a simple click of buttons and cascading screens:

> Interactive multimedia has to be more than just software that you click on to bring up a different pop-up or text-menu. 'Interactive' has to mean more than point and click - it should be involving and personal. It all comes down to concepts. A brilliant idea that works interactively ... is a way that makes sense, and that makes it a more appropriate tool than a book or a video or a set of crayons.

Moore (1989) discusses three forms of interaction that are also relevant to distinguish when thinking about technology-mediated learning. The first two relate to learner-learner and learner-teacher relations, and technology is only a mediator of interaction, serving as a means for individuals to exchange information and build knowledge on each other. The third form of interaction described by the author, and relevant to this research, refers to the relationship between learner and content. This interaction occurs between the individual and the study material, which is in a virtual format and is embodied in electronic equipment, such as a computer or a similar device. In line with the above, Wagner (1997, p. 20) explains that "Interactivity […] seems to emerge from descriptions of technological capability for establishing connections from point to point (or from point to multiple points) in real time." In this way, we consider that what Moore (1989) calls learner-content interaction in his work, defines the same idea of interactivity from the panorama presented in this section.

When the learner works directly with the virtual content without contact with other learners or teachers, interactivity plays a fundamental role. It is necessary that interactivity provides an engagement in which the content awakens connections in the learners' mind, propitiating learning moments. Yet in Moore's words, "without it, there cannot be education since it is the process of intellectually interacting with content that results in changes in the learner's understanding, the learner's perspective, or the cognitive structures of the learner's mind." It is worth noting that Moore's assertion meets the idea of learning as an internal event to the individual, in which there is a processing of information in the contact with the object, giving rise to new knowledge. Thus, it is paramount to take into account the cognitive aspect of the individual when interactivity is desired to lead to successful learning.

In consonance with this panorama, the study of Downes and MacMillan (2000) brought, applying an interview to interactive media users, the characteristics that the participants most related to the concept of interactivity. According to the authors, the ideas of control, responsiveness, and perceived goals were highlighted in the research. Regarding the aspect of control, the question of being able to move in the virtual space with a certain autonomy and the possibility of decision-making is positive. With regard to responsiveness, this feature is directly linked to the possibility of feedback the system offers, which increases its level of communication with the user, making it consequently more attractive and engaging.

Sims (2003) also investigated which features users of virtual learning environments (henceforth VLE) relate to interactivity. The author analyzed answers to three questions, which are respectively: "What do you see as the greatest benefits of interactivity in the learning process?; "What do you see as the main characteristics of interactivity?"; "What makes an educational multimedia product interactive?" The participants were 68 students of an undergraduate course in interactive and multimedia learning from an Australian university. The students' responses were grouped according to the following topics: engagement, communication, control, design, individualization, and learning. The study found that users with more VLE experience identified the communication aspect as the main characteristic of interactivity. On the other hand, users with less VLE experience determined the control aspect as an inherent trait of interactivity.

The results evidenced by the work of Downes and MacMillan (2000) and Sims (2003) lead to the reflection that having the possibility to make decisions about what

to do in the virtual environment – an aspect linked to the concept of control – is essential for content to be explored by users according to their needs and learning styles. Thus, with the free user to walk a less rigid path through the virtual material, there is a sense of autonomy and personalization in the process, which can increase user motivation and engagement, factors that are also important for successful learning in VLE. Koolstra and Boss (2009, p. 379) also show that the control feature stands out when discussing what constitutes interactivity. In the words of the authors, interactivity is "the degree to which two or more communication parties [human or computer] act on each other in an interrelated matter."

It is worth highlighting that the most experienced study participants in VLE in Sims' (2003) research considered the communication aspect as more significant, unlike less experienced participants, who pointed to the feature of control as more important when it comes to interactivity. This finding, also present in the work of Downes and MacMillan (2000), draws attention to the need of building an environment rich in feedback to the user, because in an interactive environment there is an expectation of a certain level of communication between learners and the content being manipulated and to underestimate its importance is to relegate the virtual educational enterprise to failure.

Bell and Federman (2013) also support the view that when teaching-learning conditions are favorable, that is, when materials are rich in interactivity among other aspects, online education has great potential for effectiveness. The telematic environment, when well used, can generate positive results, as it allows the creation of unique learning experiences, sometimes not achievable in a traditional classroom. We understand, thus, that there is a challenge in finding a way in which interactivity expands from a relationship of simple stimulus-response towards a somewhat less programmed and more proactive relationship between individual and object. And for this to happen, we need to recognize technology not only as an alternative medium for interaction but the evolution of forms of communication and learning.

Although the interactive potential of a teaching-learning application is of unique relevance, we cannot ignore the interaction capacity of the subjective part of the dynamics – the learner. From the point of view of Nicol (2010, p. 503), "[…] while the quality of the feedback comments is important, the quality of the students' interaction with those comments is equally, and perhaps more important." This means that neither well-architected feedback is infallible. The paths that learners take to make sense of

feedback and their response to it is an unpredictable phenomenon. Even considering the cognitive processes that it can or should elicit from the moment of its elaboration, affective variables present when learners receive and understand feedback can interfere in taking good advantage from it (STORCH; WIGGLESWORTH, 2010). Furthermore, careful elaboration for minimally satisfactory feedback is imperative in order for the probability of desired effects to remain high.

Considering that the human-machine interaction will permeate social daily life intensively for decades to come both in formal and informal contexts, the perspective under which language learning is intersected by intelligent technology will certainly shift. Knowledge building will be fostered by programs fed with data from various parts of the world that will inevitably change their structures throughout time. Languages, and especially English, extensively transmuted by the collective, will also be transformed by artificial intelligence, which has been learning and producing language, even if still in a controlled way. In addition to subjects who change the language by use, natural language processing, in a constant process of coding, decoding, fragmenting, and analyzing, will provide other learners with another language – an artificial interlanguage.

Following this line of thought, Martins and Viana (2019, p. 515) state that

> “not only does 'humanity' need to be removed from its ontological pedestal, but the very idea of 'matter' - 'machine' or 'body' - must also be rethought; not as an end in itself, but as a (re) vigorous formation, paying attention to the constant productions of its own streaks and to its own indeterminability between the precise boundaries of what is 'human' or 'machinic'.”[10]

in order to call the attention to the need of a deconstruction of the traditional saussurean dichotomies, the speaking minds and its fixed notion of language, implying an unaltered transmission of messages which puts actants and language in a completely passive position (BLOMMAERT, 2014). Messages and language, in this sense, are manipulated and consequently mutate in a process involving all the participants of the interaction, organic and non-organic ones. Acknowledging this is to think in a post-human perspective of meaning production, considering the agency of enunciative acts can originate both from biological and artificial entities.

[10] Our translation. In the original: “não só a “humanidade” precisa ser removida de seu pedestal ontológico como a própria ideia de “matéria” – da “máquina” ou do “corpo” – precisa também ser repensada; não enquanto um fim em si mesma, mas sim enquanto uma formação (re)vigorante, com atenção para as constantes produções de suas próprias raias e para a própria indeterminabilidade entre as fronteiras precisas daquilo que é “humano” ou “maquínico”.”

Relating this prospect to language learning tutoring systems, there is currently a team of experts who program them in a supervised way. However, they are not far from having more independent, high-level learning systems available for use. By looking at what a post-human future holds, subtle changes in the interpretation of algorithms will result in linguistic corrections minimally deviant from what is known as standard, which will generate feedback that may interfere with the users' interlanguage development, which may be discretely constructed differently. This, in the long run, will engender new conceptions of what language is and means, our views of language learning and communication. By decentralizing the perception of what to be human represents, as Haraway (1991, p. 180) posits in her cyborg manifesto, we might be able to find "[…] a way out of the maze of dualisms in which we have explained our bodies and our tools to ourselves."

Understanding the mechanics and the reach of these automated systems can shed a light on how and how much we, as language teachers, can partake our praxis with automated technology and build a learning environment that benefits both teacher and students. There is a need to find a path in which the human – highly adaptable and contextualized – can ride along with the machine – still highly predictable and impartial – in order to make learning and teaching fit what 21st century educational panorama is presenting. Otherwise, these virtual tools may fall into oblivion and language-learning professionals will miss an important step towards a posthuman paradigm.

This chapter was dedicated to relating theoretical as well as practical conceptions related to the area of languages teaching and learning mediated by technology. Furthermore, we focused on identifying the contribution that this study brings to Applied Linguistics concerning the extent to which automated teaching tools can offer assistance and thus compose an instructional scenario that incorporates the best of the human as well as the digital world in order to reach more sufficing learning outcomes. We also explored themes that are relevant to the composition of the framework necessary for the analysis of the data and that base our theoretical proposition. The following chapter will be devoted to describing the methodology chosen to account for the collection and generation of data from this research.

5 Research methodology

In this chapter, we will discuss the methodology used to carry out this investigation. Firstly, the proposed objectives will be resumed. Next, the context in which the research was developed will be characterized. Later, the software selected for the study will be described in depth and the procedures used for data collection and analysis will be specified.

Based on the general objective of this research, to verify the functioning of the application *Write and Improve* as an automated tool for evaluating writing in English, a set of aspects related to the capacity of interactivity and the provision of automatic feedback were determined, Among the aspects concerning automatic feedback – its most distinctive feature –, the specific objectives of the present investigation are intended to account for a) the types of errors that the program identified in the subjects' texts; b) the form of presentation of the feedback to the subjects; and finally c) the interactions that emerged among learners, application and teacher through *+Class View*.

In order to better address the questions presented, the qualitative methodology is considered the most favorable way, given the documental and exploratory nature of this investigation. In view of this scenario, the objective here proposed is not to find a finite answer to a given question, but to characterize a process in constant transformation.

Allied to the qualitative character, this research holds a documental character, since it uses materials not yet treated analytically for the elaboration of reflections about the subject (GIL, 1999). The documental analysis is an important technique of data collection since it complements information obtained by other techniques and also reveals new aspects of a theme or problem (LUDKE; ANDRÉ, 1986). Thus, once the object of study has not yet been researched in depth, the analysis starts during the data collection, in a simultaneous process of uncovering pre-established evidence and interconnecting theoretical concepts and hypotheses to account for them.

The exploratory nature of this work suggests a post-positivist view, in the sense that we do not seek a single truth once we consider the world open to multiple interpretations. To account for what is in the world, an intuitive methodology is a relevant path when working without a predetermined hypothesis. Knowing the world

with the support of metaphors, following creativity and suspicion about a theme are legitimate ways of understanding it (O'LEARY, 2004).

Yet in the words of O'Leary (2004), with reference to the position of the researcher, unlike an idea of neutrality long pursued to promote credibility, he or she is involved in the research, clearly admitting the influence of his or her subjectivity on the results he or she finds. Thus, the results are tied to a more specific context that may not always be generalized. However, this does not mean that they are not reproducible in another scenario.

The aspects listed at the beginning of this chapter will be analyzed under the concepts addressed in the theoretical basis of this work, that is, in the perspective of human-machine relations in language learning, which embrace the question of interactivity as an instrument that enables learners' interaction the with the program in a useful way. As respect to the relationship between subject and object in the digital environment, the configuration of automatic feedback can contribute to different ways of learning to take place. Different interactions – from the so-called conventional classroom situations – are manifested in this scenario and may trigger the use of differentiated communication strategies by learners, which may ultimately point to a redesign of FL's writing learning path mediated by this type of automated online tool.

The following is a description of the research context and the subjects involved, the description of the program object of this study and, finally, the presentation of the procedures for collecting and analyzing the data generated through the program.

5.1. Research context

In order to be able to account for the research data, a methodological framework was sought in the literature of the area to meet the exploratory and documental view of this investigation. However, the evaluation of automated online tools for the improvement of FL writing is a field in expansion. Therefore, the analysis of these types of data become a process of discovery, in which new information is found and mechanisms for better examining them are developed along the way.

The evaluation of the efficiency of computer programs and web pages aimed at language teaching and learning has as main objective to supply the teacher/tutor/instructor(s) with pertinent information to their practice so that they can decide which tool is most appropriate to implement in their context of action. In general,

in order to comply with this objective, the teacher/tutor/instructor(s) use a previously prepared list where they examine items in order to conclude whether the tool in question is in accordance with their pedagogical need. The common items in this type of list are related to educational aspects, according to Davies (2012, online), such as:

> Is the level of language that the program offers clearly indicated?
> Is the learner offered useful feedback if s/he gets something wrong?
> Can the learner seek help, e.g. on grammar, vocabulary, pronunciation, cultural content?

In addition to these issues, there is also a concern for more technical items such as ease of navigation, frequency of content update, the fluidity of the interface, need for plug-ins to view material, the functionality of links offered, among others.

The purpose of this research goes beyond the judgment of the appropriateness of a learning object for use in a particular instructional context. We propose to bring relevant questions regarding the feedback and interactivity present in the tool and how these elements act in the process of assisting the development of FL writing skills under a post-human perspective of human-machine interactions in the 21st century. To do so, we will use data from the program itself, captured by the teacher-researcher's own observations, the texts produced by the subjects, and field notes made during the experiment with the subjects interacting with the tutoring system.

In view of the proximity of the researcher with potential subjects to carry out the research, the private English Language course where the researcher is a current teacher in the south of Rio Grande do Sul was a favorable environment to implement the use of *Write and Improve*. The course has an intensive as well as an extensive study program, consisting of five modules, ranging from basic level to advanced intermediate, according to CEFR parameters.

In the intensive program, the course includes three initial modules that last three months each and the material is digital and gamified. For each of these modules, there is a digital book containing ten lessons organized in a fixed sequence in the flipped classroom model, where the students work with the digital material before attending a face-to-face lesson about the content. The fourth module has a physical book with a duration of three months with a weekly class of three hours. The fifth module lasts six months and has also a physical book and a three-hour lesson per week. The course lasts a year and a half in total.

The attendance of the classes is weekly, and depending on each module, the classes go from one to three hours of duration a week. The first module has one hour per week of classroom instruction; the second module has two classes and the third module has three hours each week. The content of the lessons made before the face-to-face lessons approach all the four language skills; there are exercises in pronunciation, grammatical structure, reading, writing and listening. The same way as the material, the lessons are taught in the target language only.

In the extensive program, the modules and the materials are the same, presenting the difference of the weekly workload and the total duration of the course. Each module lasts one semester and the weekly workload is fixed in two hours. The course lasts a total of three years.

5.2. Participants

The subjects who participated in the research consisted of three students, two females and one male. Participants were invited to test a new tool to assist their writing production in English during the second semester of 2018. All participants were part of a private language course where the researcher worked during the conduction of this investigation. The researcher was the teacher of the participants before and during the period in which they worked with the tool. Of the three students, two were beginners and one had just completed the course. The subjects' age range was between 18 and 25 years. The three subjects signed a consent form and were aware of the experiment procedures, being able to contact the researcher or quit the study at any time. (Appendix A).

Table 3 - Subjects' age and proficiency levels when data collection started

Subjects	Proficiency level	Age
Jéssica	Beginner	21
Jaqueline	Beginner	24
Daniel	Intermediate	18

5.2.1. Beginner-level participants

The beginner participants had a frequency of two weekly meetings for English language classes lasting one hour each. They had already had one semester of classes prior to the start of the experiment with the online tool. In addition to the classes, they had a group in the *WhatsApp* chat application to communicate and mainly to have access to the teacher-researcher in moments outside the classroom. The creation of this group was done by one of the members of the class the participants were part.

As for the invitation to these beginner-level learners, the motivation of the choice was due to their level of commitment to learning that had been observed in the semester prior to the experiment as well as the good relationship between teacher and students. In this way, we believed that engagement with the proposal could happen more fluidly and would thus be fruitful to both sides. The beginner-level participants were studying in the second module of the extensive program, having completed a total of two semesters of classes, and the second semester using the *Write and Improve* program to practice their written skills.

5.2.2. Intermediate level participant

The third participant was a former student, thus he had no weekly contact for formal language instruction during the experiment and met with the teacher-researcher sporadically when both judged necessary. This former student participant had had one and a half years of English classes prior to the beginning of the use of the online tool. The communication was held mostly via *Messenger* chat application, as they were friends on the social network *Facebook*.

It is worth clarifying in this section the choice of the former student as a participant in the research and information on his motivation to take part in it. Regarding the researcher's choice, her interest was to search for individuals studying at the school where she worked who had a level of proficiency above beginner level in order to collect data from learners from different proficiency levels. In addition, among the students contacted, this subject had an intrinsic motivation to invest time in improving his writing skill, since he had the intention to apply for vacancies in North American universities in the year following the experiment. Thus, the subject noticed in the

invitation a way of practicing his writing to produce good essays, which are required for the evaluation of candidacies for the vacancies of the desired universities. From this personal desire of the participant, the researcher observed a possible greater bond to the proposal and consequently a richer experiment as a result. The subject had already finished the course by the time the experiment was applied. He had entered the first module of the intensive study program available at the school, having completed a total of 174 hours of English classes.

5.3. Data collection method

Data collection began in February 2018, starting with the response to an email sent by the researcher to the *Write and Improve* support team. From that moment on, regular contact and exchange of material and information were established with its representatives. The researcher began to make daily use of the program in order to collect any useful data to the description that was proposed. The support team shared scientific articles aimed at the development of its database and was open and interested in the results that would be found in view of the recent product launch.

In August 2018, this approximation resulted in an offer to a three-month free trial for the premium utility of the product, called *+Class View*. This *Write and Improve* service is designed for use with student groups, having a per-student cost. The *+Class View* proposal aims to offer the teacher/tutor/instructor(s) a platform to work the writing skill in English more independently with their students and to have the aid of the automated feedback to correct texts in the surface level, identifying and/or correcting inappropriate words and phrases. Thus, the teacher/tutor/instructor(s) have more time to provide more personalized feedback, taking into account who writes the text and the language teaching-learning context in which they are inserted. Next, the space reserved for the teacher/tutor/instructor for the provision of individualized feedback is highlighted (Figure 4).

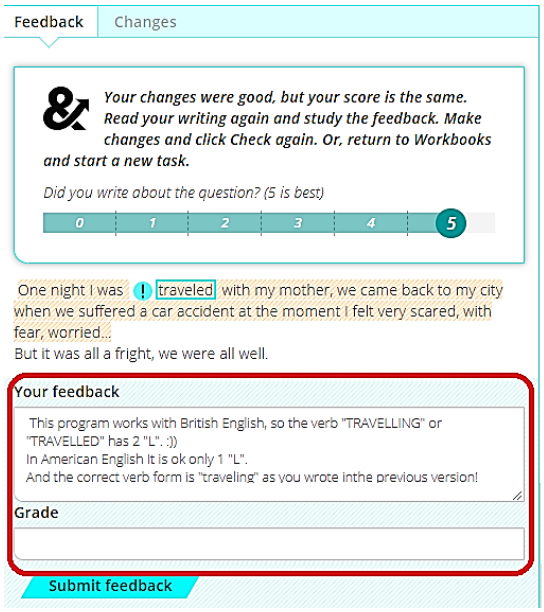


Figure 8 - Area reserved for the provision of human feedback

Having this tool available, the teacher-researcher created workbooks for the subjects and applied the experiment during the second semester of 2018. In the workbook, there were 10 writing tasks, accompanied by an illustrative image, a topic suggestion, and a word limit. The tasks were released weekly and the teacher-researcher kept the subjects informed of the new tasks through *Whatsapp* or *Messenger* chat applications. The subjects did not have a deadline to produce the texts but were informed that the textual production was part of the syllabus content of the semester and would be evaluated.

The beginner-level students had a workbook distinct from the subject at an intermediate level. Throughout the week, one of the classes of the beginner students was devoted to the practice of oral skill – called conversation class. Thus, this conversation class always had a theme to be worked on and the writing task for on *Write and Improve* was based on this theme. This strategy was thought to take into account the fact that the subjects would easily create a text about a topic after it was debated in class, helping in the construction of ideas, the emergence of appropriate vocabulary and clarification of doubts. Due to the weekly contact with the participants, the teacher-researcher used to comment on the topics that were already available in the workbook and recommended the subjects to practice writing regularly.

In order for subjects to have access to the workbooks planned by the teacher-researcher, there was a need for a previous registration. In this way, it was necessary to carry out the registration in class in the presence of each of the participants, with the creation of a user name and a password for the use of *+Class View*.

Next, a screenshot of the subjects' desktop in one of the tasks (Figure 9) is presented and in the sequence the list of tasks made available to the subjects of the beginner level, in the respective order in which they were launched.

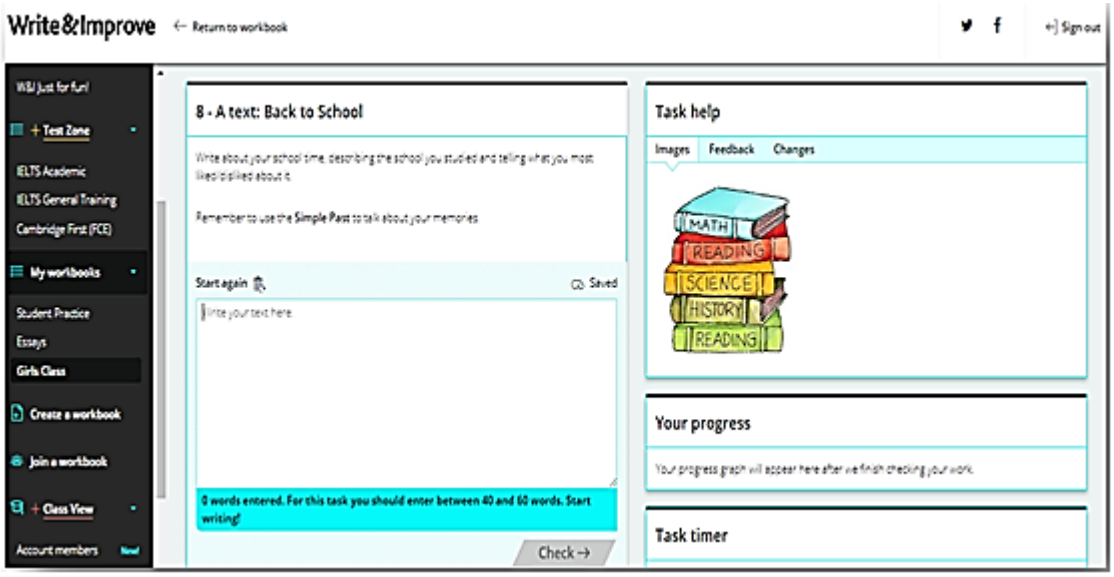


Figure 9 - Learner's desktop for writing practice

In order to observe the written production of the intermediate participant, a new workbook was created, with tasks that were relevant to the goal of writing texts for an American university application. Before the creation of the workbook, participant and researcher discussed topics that would be useful and interesting to write about. Ten tasks were released. The subject could choose the ones he would write and there was no deadline for the tasks. As the texts were produced, the subject would contact the researcher via *Messenger* chat application to inform her about the task and also to comment on the challenges of the task and mainly about the feedback of the program.

From the 10 tasks written on *Write and Improve* in each of the two workbooks, not all were used by the participants, given the previous clarification of the non-obligation to perform all the proposed tasks. All participants were given fictitious names for preserving anonymity. The researcher selected the themes of the tasks for the beginner participants and some themes made available to the subject at the

intermediate level were chosen under mutual agreement. The texts required between 30 and 250 words. The topics, their complexity, and extent were determined according to the level of proficiency of the participants. Next, the topics that the participants chose to produce their texts are presented, respecting the division of the level of the learners, as introduced previously.

Table 3 shows the relation of the number of revisions made in the texts produced by the beginner level subjects:

Table 4 - Number of texts rewritings according to the topic

Themes ⬇ Subjects ➡	Jaqueline	Jéssica
Emmergency Situations	2	1
Describe a romantic, funny or scary event in your life.	2	4
Describing your hometown	2	2
Little money, much fun	11	1
Have you ever...?	5	5
Party animal or party pooper?	2	3
Relationships		5
Back to School		3
Trip advice to a friend		6
Life goals and dreams		7

Table 4 shows the relation of the number of revisions made in the texts produced by the intermediate level subject:

Table 5 - Number of texts rewritings according to the topic

Themes	Daniel
An essay: An autobiographical essay: A difficult decision.	6
An article for a magazine: Free time interests.	9
A letter to a newspaper: Banning traffic from the city center.	9
An opinion essay: A long calm life or a short intense life?	9
A competition entry: Important things to know about me.	5
An answer to a question: Tufts.	41
A descriptive essay: A character in a film.	
A letter: Living in a different country.	12
An article: Mobile phones.	

Therefore, 24 texts and all their corrected versions comprise the corpus of analysis of this investigation combined with *Write and Improve* itself, also under scrutiny. The analysis will also be complemented by messages exchanged during the

experiment via chat applications between the teacher-researcher and the participating subjects and field notes of the classes delivered face-to-face.

5.4. Data Analysis procedures

Starting from the descriptive focus of this investigation, the analysis will be based on the types of automated feedback that the versions of the texts received, investigating the textual transformations and the level of proficiency fluctuations fomented by it, crossing the data with the dates and times in which the rewritings were performed. The aim is to understand the interaction between the subjects and the program through these different types of feedback that the system makes available, verifying the effectiveness in helping learners to employ communication strategies favorable to the correction of their errors in an autonomous context.

In order to organize the analysis process, we sought to systematize the feedback models available in the program. However, it is necessary that they are generated from the error committed and identified by the algorithm. In this way, 10 different types of feedback emerged and were collected from the texts produced by the subjects. Here are the feedback types found (Figure 10):

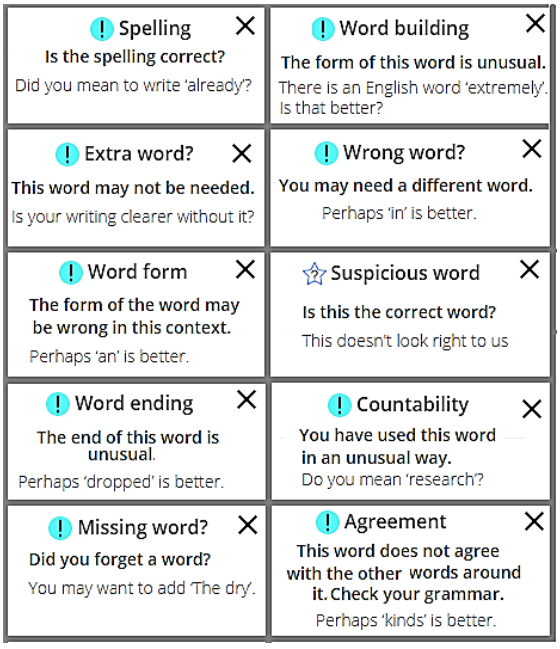


Figure 10 - Feedback types on the word level

From this organization of the types of feedback illustrated in figure 3, we will discuss how it was presented, considering the syntactic context in which it occurred, the message with the correction of the error and the communicative strategies triggered or not by the automatic feedback. These considerations will also be crossed by the analysis of the level of proficiency fluctuations of the texts during the corrections made by the subjects and the interaction of the subjects with the teacher via online tool and the implications of all these coexisting factors in the improvement of their writing.

In order to account for the impact of automated feedback in the rewriting of texts, the taxonomy of communication strategies proposed by Dornyei and Scott (1997) will be used as a reference. From the 33 categories described by the authors, only 13 can be observed in writing in contrast with oral production, in which their taxonomy was based on. They can be divided into two groups, one that the strategy is a reflection of learners' risk-avoidance, such as abandonment, reduction, omission, and substitution. The other group relates to CSs that reflect learners' risk-taking in conveying the intended message: paraphrase, approximation, restructuring, literal translation, foreignizing, word coinage, code switch and all-purpose words. The texts will be examined in order to delimit which of these strategies were directly raised by the interaction with the program, besides pointing out the mental processes possibly involved in this practice through the communication between the subject and the automated system. In addition, the field notes made by the researcher during the experiment and the messages exchanged with the subjects in online chat applications (*Messenger* and *Whatsapp*) will serve as a complementation in the search for establishing the influence of human-machine interaction on the improvement of learners' writing skill using the program *Write and Improve* modality *+Class View*.

In the following chapter, the analysis and discussion of the collected data will be presented, seeking to bring to light the process of improvement of the FL writing skill fostered by an automatic evaluation program considering the relations that the interactivity and feedback between subject and intelligent system make possible.

6 Findings and discussion

This chapter is divided into four sections. The first section brings the history and the description of the application *Write and Improve*. The second section presents the proposal of the modality *+Class View*, which was the object of the experiment with this research's subjects. The third section shows the configuration and categories of the automated feedback present in the program. Lastly, the fourth section brings the triad of learner-teacher-program interaction, evidencing its implications for the learning of FL writing skills autonomously.

It is relevant to mention that the organization of the data analysis was based on the participants and not on the feedback types previously presented. We decided to look at the phenomena under this angle once we believe that the starting point of the interactions is the subject. All the events that follow are primarily a consequence of the human input into the system, that is, the text the subject writes in the application. Therefore, section 4 of this chapter is divided into 3 sub-sections, respectively: 4.1) Daniel, the intermediate learner; 4.2) Jaqueline, a beginner learner; and finally 4.3) Jéssica, also a beginner learner.

6.1. History and description of *Write and Improve*

Write and Improve is a free automated cloud tool for improving English writing for non-native speakers at different levels of proficiency. It is a pedagogical tool that provides a VLE for the production of texts with automatic feedback in several levels in an interactive way. It was developed by the English University of Cambridge in partnership with the company ELIT (English Language ITutoring) and the institute ALTA (Institute for Automated Language Teaching and Assessment), among other companies, being this its first product. In 2014 *Write and Improve* was released as a beta version and in 2016 started operating in its final version. Its use can be as a complement to face-to-face classes as well as for an autonomous study via smartphone, tablet or computer.

Assuming that feedback is an essential part of good development in a foreign language, *Write and Improve* provides immediate and automatic feedback for learners with an interest in being proficient in English. Its focus is Cambridge University proficiency tests and the IELTS (International English Language Testing System)

General English or Academic English. There are 250 tasks available in the application, divided into four workbooks at beginner, intermediate, advanced levels and just for fun, with less serious themes, as well as two separate workbooks – one for IELTS tests and one for Cambridge English First, the former FCE (Figure 11).

Write and Improve also features a message box for the users to identify their level of proficiency in order to better target it. In addition, there is the possibility of creating a profile to provide the system of personalized information that may be used for future statistics of the program itself. The questions in the demographic questionnaire are about whether the user is a learner, a teacher, or both, the primary purpose for using the program, whether they are preparing for some specific proficiency test or preparing students for it, and the main learning objective (Figure 12).

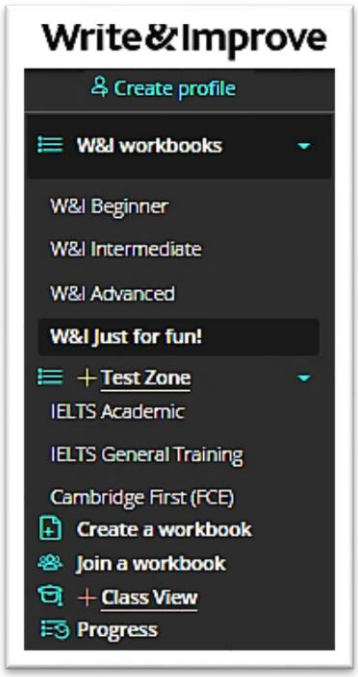


Figure 11 - Write and Improve taskbar

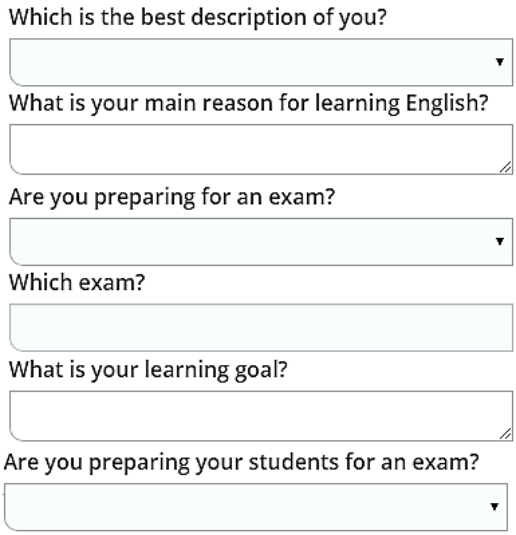


Figure 12 - Demographic questionnaire to build user profile

Write and Improve feedback is based on the Common European Framework of Reference for Languages (CEFR), a well-known framework for assessing the level of proficiency in foreign languages in a standardized way (COUNCIL OF EUROPE, 2001). The CEFR table and the Cambridge proficiency tests are closely linked, as the latter served as a parameter to establish the B2 level present in the CEFR table. Sometime later, other Cambridge exams were designed to examine the remaining levels of the framework, such as Cambridge English Proficiency for level C2, Cambridge English Advanced for level C1, Cambridge English Preliminary for level B1 and Cambridge English Key to level A2.

The Cambridge English Corpus (henceforth CLC) established the corpus that forms the basis for the program to build feedback. In 1993, the Cambridge English Language Assessment, responsible for proficiency examinations in collaboration with the Cambridge University Press, gathered texts produced by learners who performed the proficiency tests along with their data, such as written test scores, age, gender, and mother language (henceforth ML). The CLC is expanding up to this day, growing its word volume annually by two million to three million and providing important data to enhance the reach of automated feedback in English language learning. The final corpus used for *Write and Improve* consists of 2,312 texts produced by 2,312 native speakers of different mother languages. The texts have an average of 200 words each and cover all the levels present in the CEFR: A1, A2, B1, B2, C1, and C2, where A is

the most beginner level and C is the most proficient level. Table 6 summarizes the set of proposed CEFR levels in single holistic paragraphs:

Table 6 - Common Reference Levels: global scale

Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Source: Council of Europe (2001) retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale on 05 Oct 2019.

The technology that supports the tool is called supervised machine learning, an unfolding area of natural language processing, in which an algorithm is created to learn from the data it obtains and produce feedback. Initially, the algorithm, which is a set of rules or instructions that the program receives to solve a problem, was fed with 30 million words with errors annotated by CLC experts since September 2016. Currently, *Write and Improve* receives an average of 1000 texts per day, which in the

same way serve as data for the production of feedback from inferences that the algorithm makes with what it has previously learned. Because the program offers the related tasks on topics that the learners must write about, the system must have received a minimum number of responses for each theme available in order to generate relevant data for the provision of feedback. The following is an example of tasks available in the beginner level workbook (Figure 13):

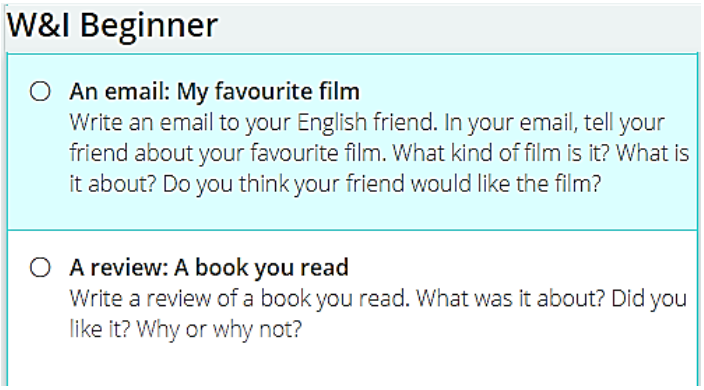


Figure 13 - Beginners' workbook tasks examples

The algorithm used in *Write and Improve* was created by Joachims (2002) called Support-Vector Machine and works as follows: once the transcriptions of the original texts are done, annotations of errors are made by a team of specialists. The second step is to separate the correct sentences from the incorrect ones and send them to the program called Robust Accurate Statistical Parsing (henceforth RASP), which does the encoding, breaking and parsing of the structures syntactically. According to iLexIR, the company that holds this technology,

> "The RASP system includes state-of-the-art modules for finding sentence boundaries, finding individual words, analyzing words to identify the word root and any suffixes, assigning part-of-speech labels to words in running text, and analyzing the grammatical relations between words and larger units within sentences. Text analyzed with RASP provides the basis for text classification on the basis of topic, sentiment, genre, reference to specific entities, the strength of specific assertions or many other facets, when combined with other open source machine learning classifiers either using supervised or semi-supervised learning techniques. The resulting annotated text collections can be indexed using open source search engines at the document, sentence or word level to provide flexible, intuitive, interactive access to text snippets and passages or to automatically create structured databases from text."[11]

[11] Description retrieved from https://www.ilexir.co.uk/rasp/index.html on Aug 28, 2019.

This process trains the algorithm that will analyze the texts written in the tool. Then, the algorithm places the texts on a scale to accommodate feedback according to the benchmarks of the tests the user is practicing. The algorithm extracts some characteristics, which are established by the responsible team and compares to a scale, in this case the CEFR, to evaluate the text. This scale is presented to the user in the form of a graph. The process described above is performed so that the users have access to the feedback called summative, that is, their texts receive a score from A1 to C2.

In general, this type of algorithm performs a classification of texts between approved and reproved for the pre-established proficiency levels. However, for the *Write and Improve* program, the algorithm was trained to perform a ranking. According to the program developers, this type of strategy is more appropriate for an evaluation based on a symbolic scale such as the CEFR, because there is no need to assign a numeric grade or limits to establish approval or disapproval.

In order to arrive at the final step of pointing out errors in the text, the rules are derived from the texts of the corpus broken into small linguistic segments, such as 1) sequences of words, 2) sequences of grammatical classes, 3) grammatical constructions, and 4) other aspects such as 4.1) sequences of unknown words, 4.2) proportion of number of words per number of sentences and number of syllables per number of words, which verifies how comprehensible the text is (Figure 14). For the system to learn and point to such segments as errors, they need to appear numerous times. Currently, the system has 80 types of errors encoded in a universe of 40 million words. According to tests carried out with the program, there are correction failures in only 10% of cases, in which the system points out correct constructions as inadequate (YUAN; BRISCOE, 2016). Compared to the evaluations of a human, the scores that the system attributes to the texts are very close, which lends credibility to the *Write and Improve* automated analysis. Note some of the aspects that the algorithm is trained to identify (Figure 15). The tagging taxonomy used to categorize the parts of speech is the CLAWS2 tagset and has around 175 different tags.[12]

[12] UCREL CLAWS2 tagset is open source and is available for download at https://www.ilexir.co.uk/rasp/index.html

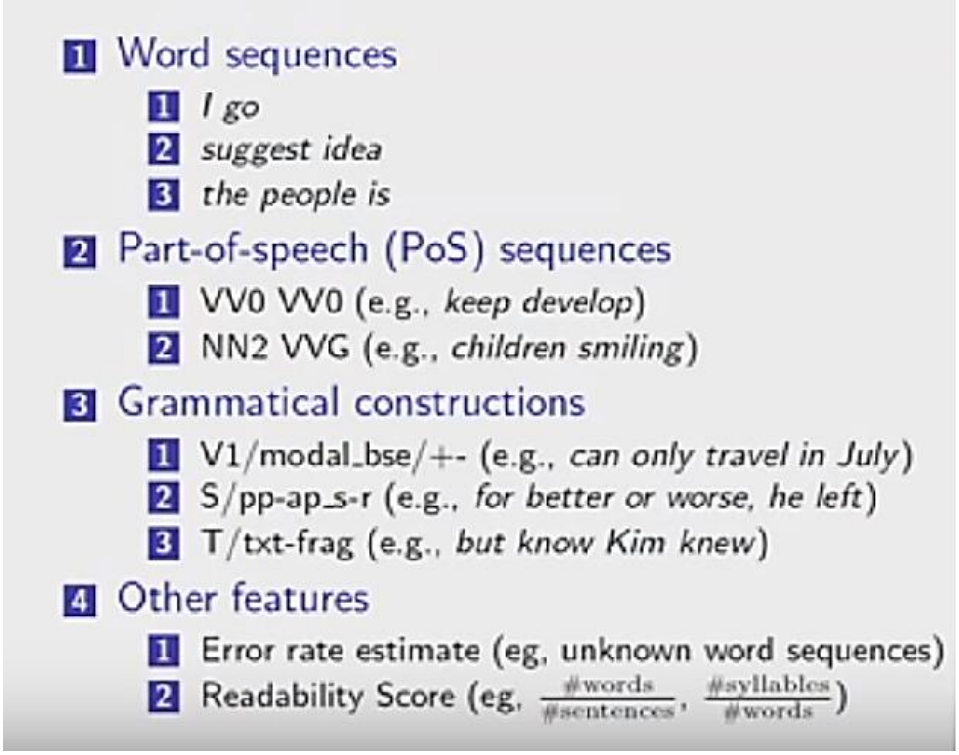


Figure 14 - Linguistic features the algorithm is trained to detect[13]

After the initial work with the RASP system, a weight is assigned for each linguistic feature, with positive values being those that increase the text's score and negative values those that decrease the text's score. The criterion of the weights is given according to the incidence of features that discriminate a text of a certain level, forming a ranking. In figure 11, below, an example of discriminatory features of a text referring to the English First test, which includes levels B1 and B2 in the CEFR scale:

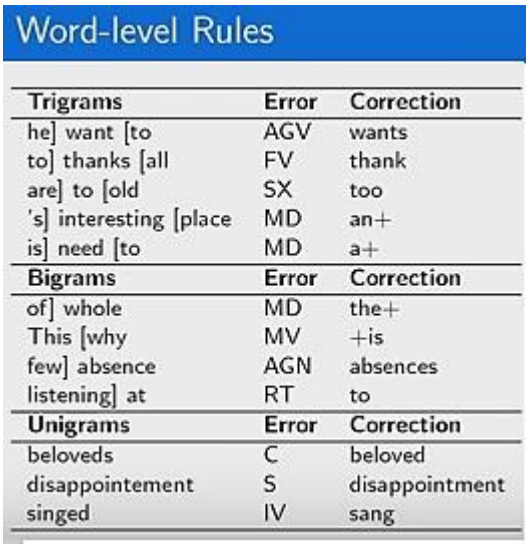
(Selected) Highly Ranked Features: FCE

Feature	Weight	Example
AP/a1	8.50278	(very) clever(ly)
T/frag	-6.12576	*But know Pat knew
NP/det_n1	6.01353	the film
VM_RR	4.26423	could clearly
PP/p1	4.26187	in the mornings
necessary	2.83127	it is necessary that
the_people	-2.73139	*the people are clever
probably	2.67593	we are probably going
P1/p_s	2.62475	after we left
VV0_VV0	-2.50546	*we go see film
NN2_VVG	2.46941	children smiling
II_VVN	-2.41127	*I want to gone

[13] Parts of Speech tags meaning: VV0 stands for base form of lexical verb; NN2 stands for plural common noun; VVG stands for -ing form of lexical verb.

Figure 15 - Discriminatory features ranking in levels B1 and B2[14]

In figure 16, we can observe how the system relates the errors in a word level and the corrections to provide feedback to learners:



Word-level Rules

Trigrams	Error	Correction
he] want [to	AGV	wants
to] thanks [all	FV	thank
are] to [old	SX	too
's] interesting [place	MD	an+
is] need [to	MD	a+
Bigrams	**Error**	**Correction**
of] whole	MD	the+
This [why	MV	+is
few] absence	AGN	absences
listening] at	RT	to
Unigrams	**Error**	**Correction**
beloveds	C	beloved
disappointement	S	disappointment
singed	IV	sang

Figure 16 - Word level errors and suggested corrections[15]

In the words of Ted Briscoe, a professor of computational linguistics and ILexIR co-founder, the initial motivation to launch a tool with this pedagogical dimension was first by estimating that by the year 2050 two billion people will be studying English and 300 million people are now preparing for language proficiency tests. This volume of people will need to practice and conduct tests during their academic and school life. Therefore, there is a concern to find ways to meet this demand, for the main objective of the university as a vehicle for education is to improve society, and one of the ways to do so is to teach English more effectively. That being said, the developer's vision is to democratize learning by providing free of charge and easy access to English learning thus enabling people to have better opportunities in their lives.

Nevertheless, not all *Write and Improve* features are free of charge. Thus, in order for teachers to use it with their groups of students at the institutional level, there is a

[14] T/frag stands for text fragment; VM_RR stands for modal auxiliary and general adverb; VV0_VV0 stands for sequence of 2 base form lexical verbs; NN2_VVG stands for plural common noun and –ing form of lexical verb; NN2_VVG stands for plural common noun and –ing form of lexical verb; II_VVN stands for preposition and past participle form of verb; NN2_VVG stands for plural common noun and -ing form of lexical verb.

[15] AGV stands for agreement of the verb; FV stands for form of the verb; S and SX stands for spelling; MD stands for determiner; AGN stands for agreement of number; C stands for conjugation; IV stands for irregular verb. The acronyms not explained here did not have their description available in the UCREL CLAW2 tagset.

fee per student to be paid. Using it as a complement to face-to-face lessons can minimize the amount of writing work correction by the teacher, which can be invested in other learning activities. However, so far, the use of automated writing evaluation systems is not a common practice in schools and universities, mostly because of teachers' pedagogical beliefs (WARSCHAUER; GRIMES, 2008).

According to the developers, the application is great for learners who are at the intermediate level (Andersen et al, 2013). At A1 level, which is the first on the scale based on the CEFR (COUNCIL OF EUROPE, 2001), the user writes very little and it is difficult to work with little written material. At the last level, C2, next to a native, there is little the system can offer in terms of tips to boost learning. If a native of the English Language, for example, writes a text in which the topic is inserted in practices for beginners, it is impossible for him to receive a C2 score, since what he wrote is within a context that includes only initial levels like A1 and A2. This means that the tool does not score the user, but rather the text. This information is available in the program webpage FAQs.

The accuracy of *Write and Improve* is 90%, that is, in 10 errors pointed out at the word level, only one is inadequate because it was not a real error, probably caused by a human error in previous error coding. This measure is given by the percentage of times that the algorithm identifies the word marked as incorrect by one teacher in 90% of the times it occurs in the data. When the system has close to 90% certainty of the error because the word was marked as incorrect 80% of the times it occurred, the program provides the "suspicious word" type of feedback. This discrepancy happens because in some cases there is no consensus among annotators about the error and consequent correction. Figures 17 and 18, in the sequence, illustrate cases where there is disagreement about error and feedback:

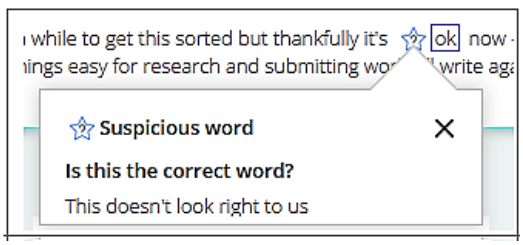


Figure 17 - Feedback fostered by disagreement in annotation

In the excerpt presented in figure 17, there is a discrepancy between the annotators involved in the coding of the errors on the writing of the term “ok”. Some consider it correct only if it is written in capital letters, so when the term appears differently written, the algorithm understands that there is the possibility of being an error but there is no consensus in the coding.

In figure 18, the determinant is pointed as suspicious and serves to alert the learner to better observe the use of the definite article in relation to the text as a whole. In this case, there is no compliance by a contextual issue and not by divergent opinions. The fact is that “the” can be used when the name that succeeds it has already been mentioned in the text. On the other hand, the term is not adequate if the noun that succeeds it is being used to convey a general idea. In this way, invariably there are codifications for both situations of the determinant in question, causing a contradiction to the algorithm, which is only able to offer feedback that is more general in order to avoid making an improper correction.

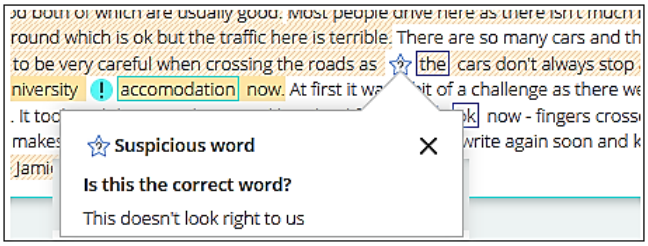


Figure 18 - Feedback triggered by ambiguity of error annotation

These context-dependent situations that permeate the text are not supplied by the algorithm since it does not yet have the ability to identify rhetorical elements such as anaphora as exemplified earlier. We can note from this case that we could easily combine paragraphs of different but grammatically correct texts and receive a high evaluation since the system does not identify if there is coherence among the ideas throughout the same text. That said, for organizational issues *Write and Improve* is not indicated, since its reach goes only to the limit of sentence, contemplating the morphological, syntactic, and semantic area. The phonetic and pragmatic area is isolated from the analysis and are not the target of feedback.

Write and Improve developers state that the language fed into the tool’s system is Learner English, once its database is built mainly from FL learners’ texts, not native speakers’ texts. However, the feedback provision is solely based on standard British

English. This can be observed in figure 19, in which the verb inflection *traveled* receives a marking and the feedback suggests the verb be written with double “l” – *travelled.* It ignores the possibility of having the verb written with a single “l”, form that is broadly used in other variants, such as the widespread American one. This certainly poses a limitation of the tool in assisting more adequately those learners whose interlanguage has been built from exposure to other variants.

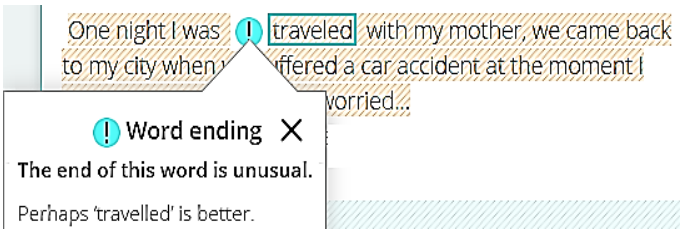


Figure 19 - Feedback on a variation of spelling

With these considerations in mind, we can consider it a very accurate program. In addition, if the scenario is a university class with around 40 students, for example, the teacher of this group would hardly point out all errors that the student had produced in his texts, by the demand that this type of task requires with this volume of productions. With this, the second motivation for the launching of this tool – besides the developers’ enthusiasm with the idea of democratization of teaching –, the reduction of the work of English teacher/tutor/instructor(s). By using *Write and Improve* to evaluate a text and point out language-level incongruities as a code, there is room for the teacher/tutor/instructor to have time to devote to feedback for problems of other origins that may occur in texts. Questions of coherence and cohesion need to be analyzed by someone who, at the time of reading, can understand the writer's intention and why certain words are used instead of others. For example, according to the developers, if there is in the text the expression “big conversation”, the system will probably point out inadequacy, since this construction is not usual of the language. However, the system is unable to infer what the writer attempted with the message, making the feedback for this something extremely complicated yet. We might think that the user wanted to say long conversation or still important conversation, but there would be a need to understand the context of the whole text for accurate feedback, which is beyond the technology that is available nowadays.

There is a study presented on the page of *Write and Improve'*s distributor, Elit, which mentions a test before its launch with 540 English language learners from 9 different ML, all at the CEFR level B1 (Andersen et al, 2013). According to the research, there were more than 3000 texts including the revisions, totaling 600,000 words. The texts had an average of 200 words each. From this data, an average of 3.2 submissions for revisions per user was verified. The highest number of revisions registered was 54. In 80% of the cases of users who submitted for revision, the level of text proficiency between the first and last submission increased. This result, then, suggests that the tips that the tool offers actually afford learning as well as present a good level of correlation with a human assessor.

The program webpage offers English and Spanish language options, but corrections are made for productions written in English only. The texts are visualized only by the learner and there is a record of all the productions made and the respective feedback, creating a history of the process of individualized learning and always open to new submissions for revision. The program developers assume that, in order to learn to write well in another language, it is necessary to be engaged in a process of practice and feedback. The program is said to be context-sensitive as a means to assess whether what the students wrote is in line with the theme which they chose to write. It is gamified with the purpose of positively dealing with users' motivation. There are 13 different badges, received according to the use of the program and completion of the proposed tasks. To mention some examples, if a text improves by three revisions of the same task, the users receive a medal that is stored in the section called "My activity & awards" positioned in the Progress tab (Figure 20). When users reach a certain number of tasks performed, they also receive awards (Figure 21). The following are examples of the prizes mentioned.

Figure 20 - Badge for improving the level after three consecutive revisions

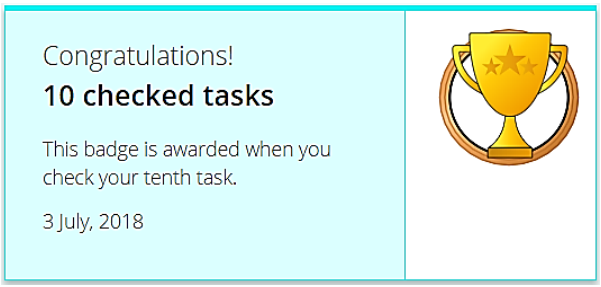


Figure 21 - 10 checked tasks award

Write and Improve also offers a modality called +*Test Zone*, which provides content targeted to those users who aim to prepare for proficiency tests such as IELTS academic, with 52 tasks available and IELTS general training, with 40 tasks available. There is also Cambridge English First, the former First Certificate of English (FCE), with 60 tasks available, which are gathered in a separate workbook. However, only one task in each of the two test models can be performed experimentally. If users wish to perform more tasks, there is a need for a paid registration. The type of evaluation performed in these workbooks is part of the same database previously described and uses the same process to provide automated feedback for the texts. The main differences between the tasks presented in the free activities are, firstly, that in the +*Test Zone* the questions resemble a little more to those that usually appear in the respective tests. The peculiarity about this *Write and Improve* modality is that you can send invitations to friends so they can know and use +*Test Zone*. That way, if someone becomes a user, the user who sent the invitation receives a discount code on the monthly amount charged, which is 9.99 euros. In addition, the user can set up his/her own tasks to practice with the themes he/she wants. The +*Test Zone* provides the creation of 143 new tasks in case the user wants to build his/her own workbook. The following is a screenshot of an IELTS Academic task (Figure 22).

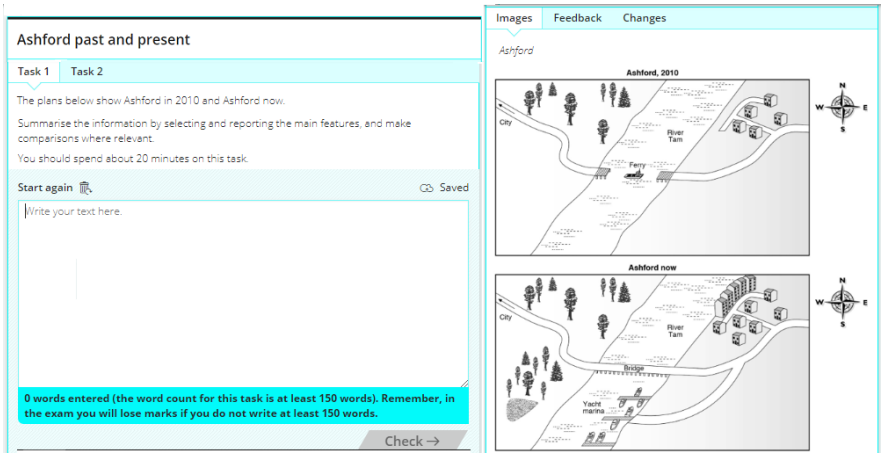


Figure 22 - Desktop of a task in +*Test Zone*

Taking into account all the tools that *Write and Improve* provides online, we can say that there are modalities available for various types of audience. There are options for users who can only use the free workbooks, and will still have access to a robust range of tasks from beginner level to advanced level. There is the alternative of a more specific material for users who are able to invest a monthly fee and have a service focused on preparing for proficiency tests in an autonomous way. Finally, there is the option of content aimed at those who teach English and want to invest in a digital application that allows a follow-up of the writing development of their students.

On this last service, thought through the view of those who teach, *Write and Improve* has developed the *+Class View*, a space for those language teaching professionals of who wish to innovate in their FL writing classes. In the following section, we will discuss the proposal of this utility that aims to encourage the use of digital tools to compose the learning process of groups of students, complementing their face-to-face studies.

6.2. *+Class View*

+Class View is a *Write and Improve* tool designed to support teachers/tutors/ instructors with their student groups. It is paid and the cost varies according to the number of students and can be implemented at any class size. It has the purpose of facilitating the work with the development of the written ability in the English Language

with regard to the creation of tasks for the practice of writing both in the classroom and at home, encouraging autonomous study.

This utility provides a framework for monitoring and controlling the development of learner proficiency. There are a number of differentiated resources to account for students' access to the program and for the completion of written assignments. The first is the possibility of creating a specific workbook for each group or, if desired, there may be a workbook for each student individually. They are arranged in the left part of the work area under My Workbooks tab (Figure 23).

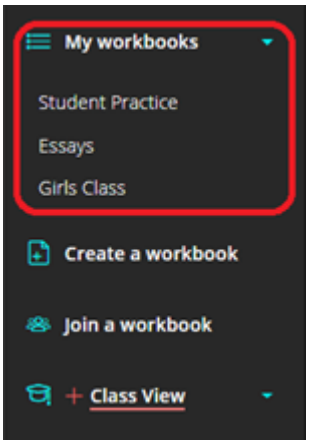


Figure 23 - Taskbar showing the workbooks created by the teacher

The teacher/tutor/instructor needs to register students under a user name and create a password for each one of them. After the first access, students can modify their passwords as they wish and there is no need to register an email address to use the application (Figure 24). After the students access it, they are confirmed as members of the class and have access to a specific workbook.

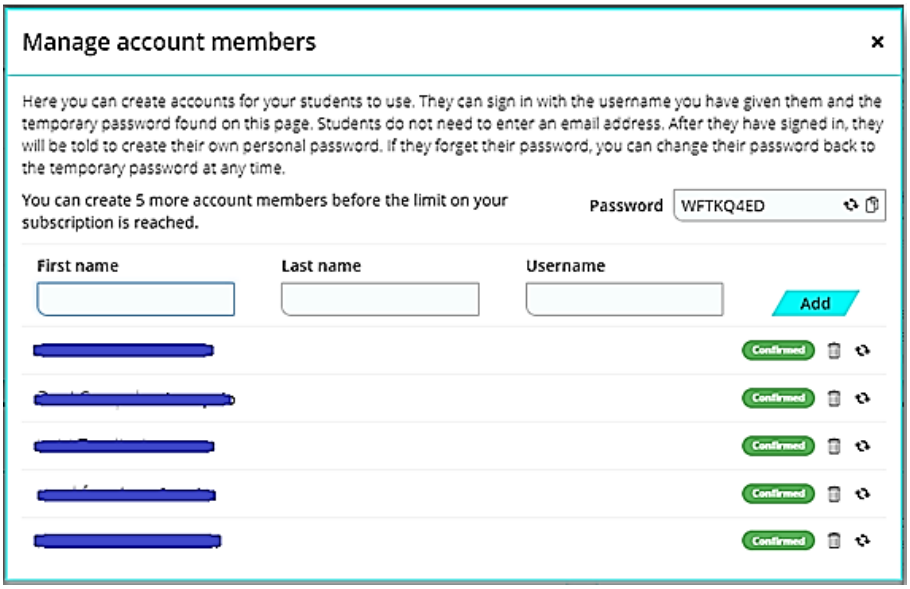


Figure 24 - Students' registration area

Once confirmed as members of a specific workbook, students have access to the tasks that the teacher/tutor/instructor will include. There are two ways for students to find out when new tasks are available: a) more autonomously, as they enter the activity book and see the signs in the shape of a small blue circle in the upper corner of the task titles and b) by email, when the teacher/tutor/instructor adds the students' emails in the workbook and clicks "send invitation" so that they can access it and check which tasks have been published.

Another way to track students' development is via reports, which can be copied to a computer in .xlsx format (worksheet compatible with the Microsoft Windows Excel program) (Figure 25). In the program, texts are stored under each student's name with date and time records of all versions submitted for feedback (Figure 26). This way, the teacher/tutor/instructor is able to carry out an evaluation over a period and may take into account other variables, such as content addressed in the classroom, student profile, etc., to analyze their progress. In figure 26, for example, we can observe that the learner took up the same text on different days as well as at different times on the same day. We should also notice that this process is partly motivated by the level of proficiency shown for each revision. Thus, the learner makes a correction and checks whether the adjustments have modified the level of his text for better or not. Just as it can be motivating, such a practice can also lead to frustration, since the level of the

text can drop between one check and another, a fact visible on the right corner of the report under the task title “Little Money” in figure 25 as well as in the revisions record in figure 26.

texts / students	Emmergency Situati		Describe a romantic,		Describing your hon		Little money, much	
	Checks	Score	Checks	Score	Checks	Score	Checks	Score
[redacted]	1	A1	4	A1	2	A2	1	A2
[redacted]	2	A1	1	A1	7	A1	1	A2
[redacted]	2	A1	2	A2	2	A2	11	A1
[redacted]	3	A2	5	A2	3	A2	2	A1

Figure 25 - Report on the number of revisions and respective proficiency level

Task: Living in a different country			
Student	Check	Date	Level
[redacted]	12	19-11-2018 18:45	C1
[redacted]	11	19-11-2018 18:44	B2
[redacted]	10	19-11-2018 18:43	C1
[redacted]	9	19-11-2018 18:43	C1
[redacted]	8	08-11-2018 17:16	B2
[redacted]	7	08-11-2018 17:13	B2
[redacted]	6	08-11-2018 17:12	B2
[redacted]	5	08-11-2018 17:09	B2
[redacted]	4	08-11-2018 17:08	B2
[redacted]	3	07-11-2018 18:31	B2

Figure 26 - Student’s revisions report on a task

In addition to these tools, there is the tab entitled *Insights*, which shows the score oscillation of students' writing. The first map is called *score range* and it allows the visualization of the fluctuation of students’ proficiency level individually and as a group (Figure 27).

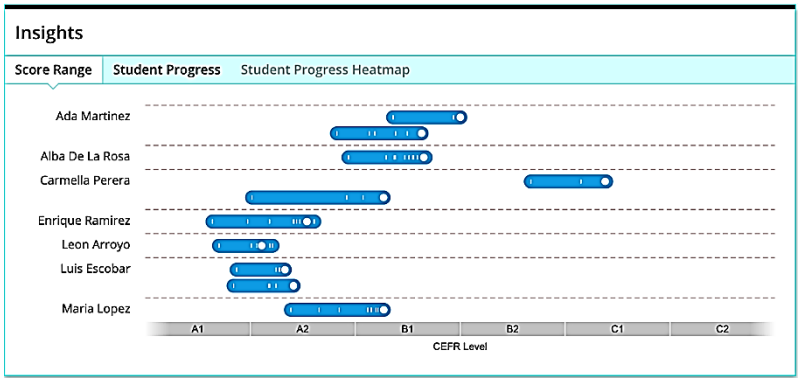


Figure 27 - Example of score range of a group of students

The second complement is called *Student Progress*, which brings a graph with information about the level of students' texts individually and in groups. In this way, it is possible to observe the writing development under a more global perspective, from the point of view of a class as well as of a specific learner (Figure 28).

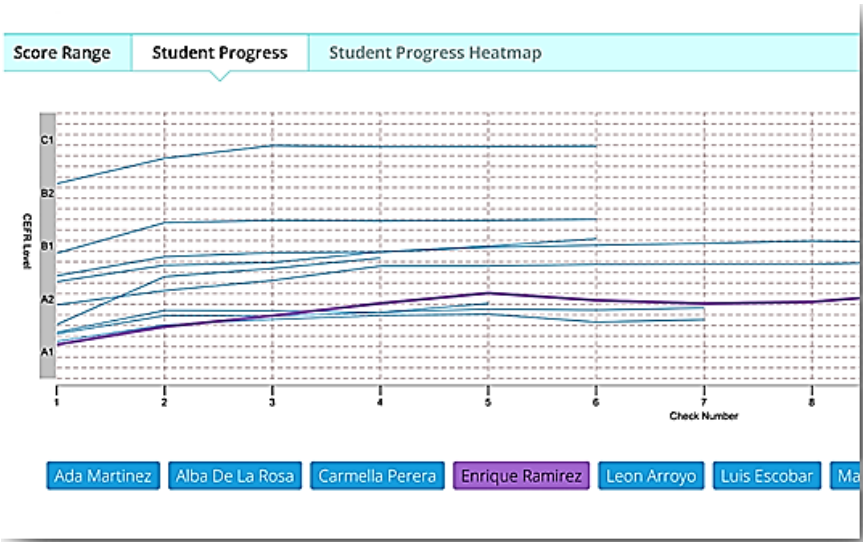


Figure 28 - Example of a writing development graph

The last supplement available to the teacher/tutor/instructor is the *Student Progress Heatmap*, which shows the evolution of text writing by means of a relationship between the number of checks and the level of proficiency achieved in each. The darker blue color circle represents the lowest level and the highest level is represented by the lighter blue color circle (Figure 29).

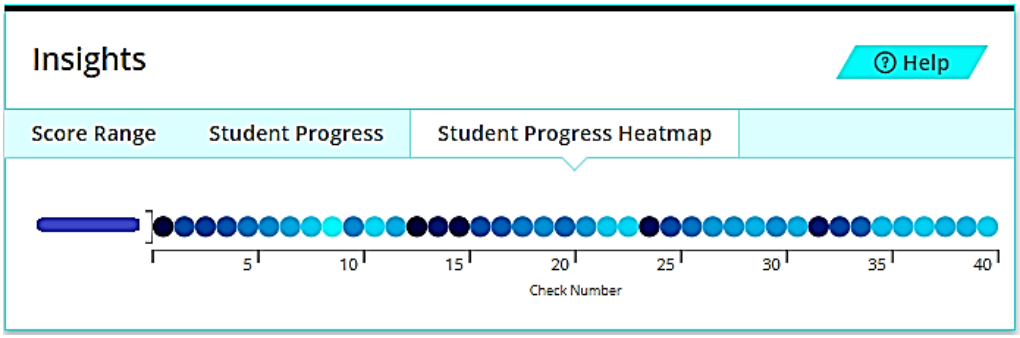


Figure 29 - Student progress heatmap according to the revision number

With these records, which provide individualized information, there is room for the teacher/tutor/instructor to think of approaches that can effectively contribute to the progress of a particular learner. Considering this possibility of differentiating the use of strategies to groups and/or students in isolation, learning is understood as adaptive, in the sense that learners can have access to differentiated materials according to their specific needs guided by their mentor. However, in using the program autonomously, outside the utility *+Class View*, that is, without access to a content created by a teacher/ tutor/instructor, this reality is not yet palpable. This means that the program itself is not yet able to select content for users and guide them along a certain path in view of the performance history of their activities. This assessment and consequent choice of content needs to be carried out by someone who is responsible for their learning, monitoring their performance through the tools that *+Class View* offers.

6.3. The automated feedback

The program *Write and Improve* takes about 15 seconds to send feedback to users, a surprising time considering the complexity and variety of the corrections it provides. However, the feedback markings and suggestions are not connected to the original text, that is, the modifications that the program recommends appear in a copy of this text, in another text box, which is not sensitive to user manipulation. So, learners need to make adjustments where they wrote or pasted the original text, looking at the marked copy. Learners are able to return to their original text by clicking the *changes* button in a taskbar located at the top of the feedback box. All the adjustment made by the learners is shown in real time on the tab opened by the *changes* command with

green markings on added words and a red color crossing out indicating removed words. So, after the desired modifications are done, learners resubmit the text for correction by clicking the check again button located right below the text box where they were making the edits. Next, a screenshot from the visualization of the text's modification (Figure 30).

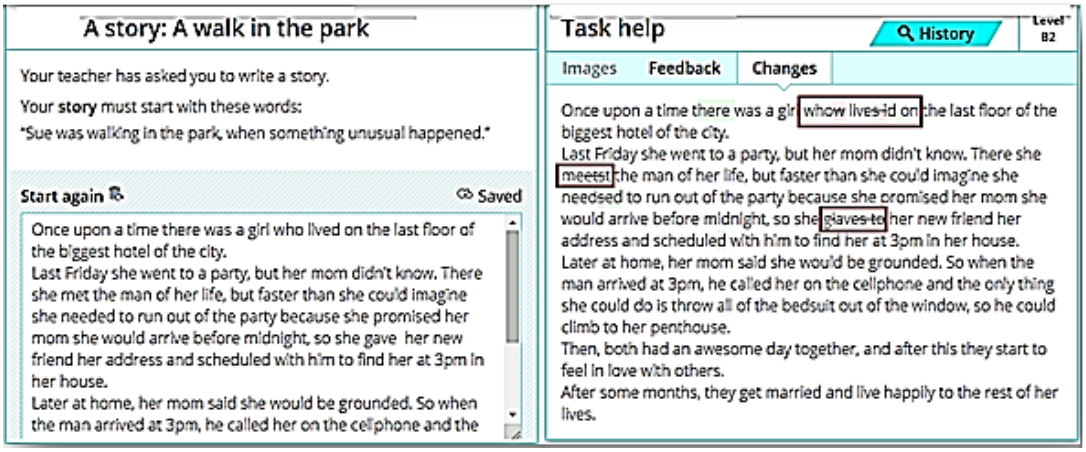


Figure 30 - Real-time adjustments visualization

At first, the design of the program can cause the impression of low dynamics because it has several tools and functions and the feedback is presented separately from the original text followed by many details. On the other hand, it is completely understandable the arrangement of all the functionalities of the program when we turn our attention to the aspect of learning as opposed to a simple proofreading mechanism. Using the application takes time precisely because it is meant to go beyond a linguistic correction, offering assistance to the learner's performance. The learners engaged in developing their writing skills in English with *Write and Improve*'s help need to work on the feedback, making reflections, rewriting terms and sentences, searching in a dictionary or grammar book. The program requires a significant level of learners' engagement so that they can fully exploit all the support the program offers to improve writing.

It is common to conceive that automated writing assessment programs should be agile in a clinical inspection of the linguistic system, but the proposal of *Write and Improve* goes in another direction, proposing a tool capable of instigating learners to study their text in order to acquire new knowledge. Its feedback configuration attempts to point out inadequacies and thereby arouse the user's curiosity in seeking solutions

to the problems encountered. An example of that is the colored marking the system applies to sentences in which elements are not very well assembled (Figure 31). The problems in the sentence construction are not pointed out in detail. Thus, learners must reread the fragments and try to figure out possible inconsistences on their own and then rewrite the sentences.

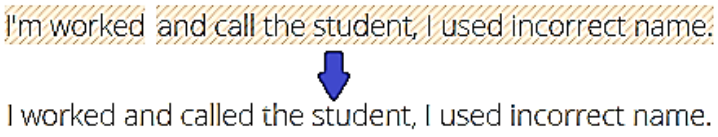


Figure 31 - Example of one way the tool instigates learners

6.3.1. Feedback symbology

The tool offers four types of feedback: a) summative, in which the text is evaluated in relation to CEFR levels; b) indirect, where there is a focus on word level, such as 1) wrong word, when there is a problem in spelling, grammar or vocabulary choice; 2) suspect word, when there is some problem in the construction of the word but the program is not able to identify it on its own; and 3) missing word, when the program identifies a gap in sentence construction. For each of these types of marking there is a specific symbology to guide learners in relation to the types of mistakes they are making.

The symbology merges geometric shapes, graphic symbols, and colors to establish communication with the user (Figure 32). A yellow triangle containing an exclamation point indicates that the following word is wrong, and a dialog box is shown below with a correction hint. A blue circle with an exclamation point inside can indicate two situations: the following word is spelled wrong and a dialog box below is displayed containing the word with the correct spelling or a message saying that the word is not required. A red triangle containing an arrow on the horizontal position pointing to the left indicates that there is a missing word in the exact position of the symbol and the next word is marked with a red frame. A pentagon in green color containing a horizontal arrow pointing to the right indicates that the learner omitted a word after the word following the symbol, and the reference word is marked with a green frame. A star in blue with a question mark inside indicates that the following word is suspect and a dialog box is displayed below with a fixed comment: "That word does not seem right to

us." This last symbol is used when the program was not able to identify the inadequacy of the word but detected an inconsistency, which will be up to the learner to recognize. See figure 32 with an illustration of the symbology described and note that the yellow color symbol is not shown:

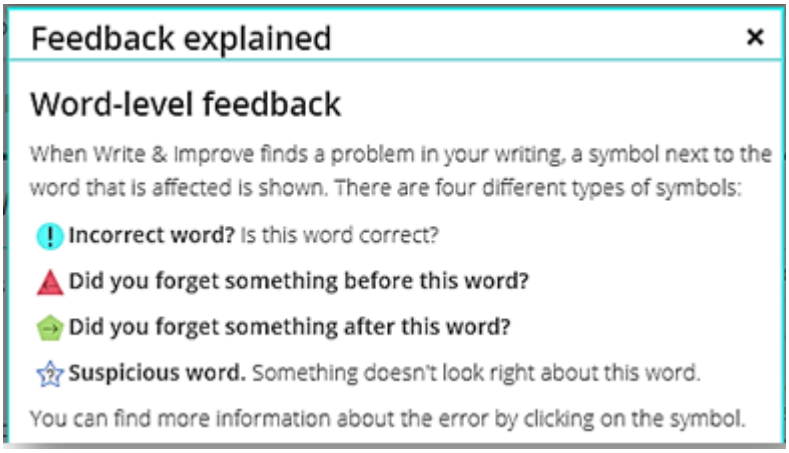


Figure 32 - Symbology created to identify the feedback hints

Taking into account the concern to create a considerably short symbology to lead learners in the writing process, there is a concern to avoid excessive marking in the text submitted to the assessment. Thus, all types of errors described above are never presented in the same check. The algorithm is trained to deliver feedback gradually, as learners make corrections and resubmit their text to revision, preventing the text from being visually polluted. According to the application developers, this guideline was established based on the view that excessive corrections become counterproductive to learners' motivation, who can evaluate their textual output negatively by visualizing so many marked inconsistencies (HYLAND; HYLAND, 2006). Furthermore, Truscott (1996), Krashen (1992), Leki (199) and VanPatten (1986) have already called the attention to the fact that grammar correction does not necessarily result in learners' interlanguage improvement because a comprehensive correction, that is, correcting every single grammar error, is too much for learners to take in at once. They also point out the time-consuming aspect of correcting all the mistakes students produce, which in the automated feedback scenario would not be a problem once it is not made by a teacher. Therefore, the authors defend a selective correction approach, which is based on learners' needs and focuses on those mistakes considered the most challenging for learners to fix.

Bearing this perspective in mind, the algorithm has a level of relevance to correct what actually compromises the text in terms of understanding. Similarly, learners are encouraged to think that there is always someone alongside them offering support. To create this effect, the messages contained in the feedback dialogs are directed at learners as an interlocutor in a conversation about the text. Thus, the messages are usually in the form of a question, they have a positive and friendly tone and are shown after each submission regardless of how much corrective feedback the text has already received.

The indirect or semi-corrective feedback is extended to sentence level, since the system highlights inaccurate sentences in two ways: 1) highlighting in solid orange coloring, indicating that the sentence contains problems in the construction that can lead to its misunderstanding and 2) highlighting in an orange stripe pattern, indicating that the sentence may still be better written, but it is understandable. When there is no marking in the construction it means that everything is correct, with no urgent need for changes. This feedback model is what stands out when comparing this tool with proofreading services, for example, as in the case of *Word* and the most recent *Grammarly*. Text proofreading software offers direct feedback, that is, they provide users with the correct answer of the inconsistencies it finds in a text. Thus, there is no commitment in terms of teaching-learning process; there are only quick solutions to linguistic problems. Proofreading software targets a diverse range of users. *Write and Improve*, however, is intended for a noticeably more specific audience – learners of English as a foreign language – and strives to offer a learning object for writing skill development. Observe the explanation that the program offers users on this type of feedback (Figure 33):

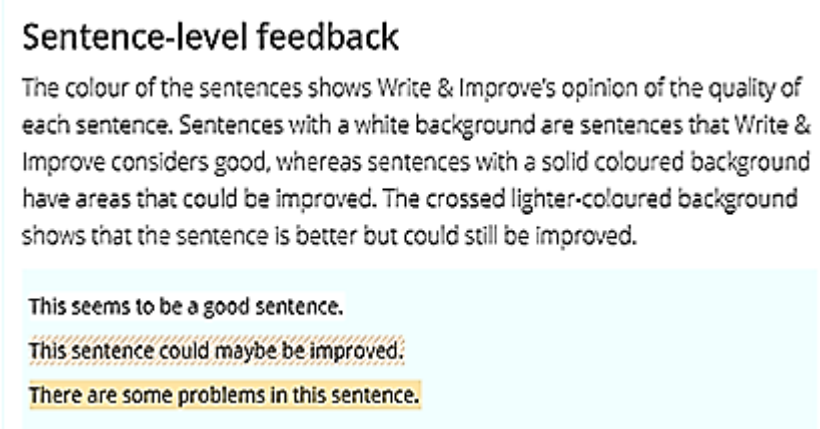


Figure 33 - Sentence-level feedback explanation

The third type of feedback is called overview (Figure 34). This feedback presentation is composed of motivational messages presented at the top right of the user's desktop after the check along with the proficiency level assessment, which is known as knowledge of results (KR). The area has plenty of functions that will allow learners to improve their text in different ways.

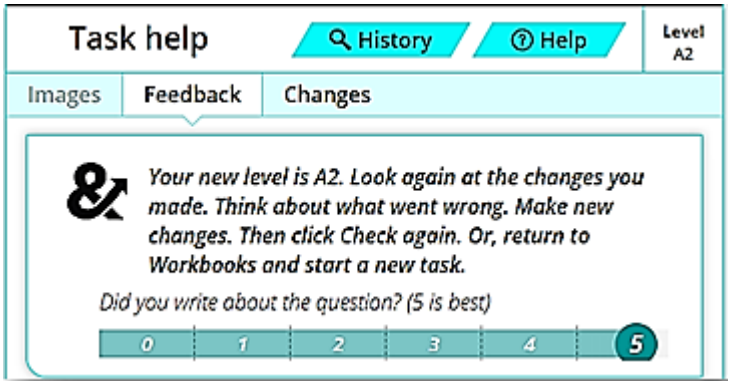


Figure 34 - Application functions taskbar

Observing the functionalities arranged in the overview feedback window on figure 34, we can see the *history* button at the top of it, a command that provides the history of checks learners did with the recording of each moment of the text. On the left side, there is the users' history of checks for the text in the figure, with a record of the task theme, the number of checks, time and proficiency level. The *help* button located at the top right of the text box contains the symbology of word and sentence level feedback described earlier to support learners on how to interpret the marking on their text.

In another segment of the taskbar, there is the *images* section, in case the task has an illustrative image. Then, there is the *feedback* section, where there is a message to the learners with an overview of their performance in the text and, finally, there is the *changes* section, which shows a copy of the original text in real time with the editions being performed by learners in a very similar way to the command present in the *Word* editor called *track changes* in the *revision* section of its taskbar. Finally, below there is the motivational message along with a bar that evaluates from 1 to 5 how close to the theme the text is, 5 being the desired one since all texts are written

after choosing a topic, in which the program bases its assessment. Below this message box, there is the text with the markings of the problems the algorithm detected, called indirect or semi-corrective feedback, at word level as well as at sentence level.

In addition to these main functions, *Write and Improve* also provides a timer and a word counter just below the text box assigned to the written output in case learners need to train their writing time to perform a limited time test, a common feature in Cambridge proficiency exams that *Write and Improve* focuses on.

6.4. *+Class View* interaction triad

In this section, we will discuss the presentation of feedback to learners using the texts produced by the research subjects in the *+ClassView* utility, based on the types of feedback available in the system. We will discuss how they were configured, considering the syntactic context in which they occurred, the message with the error correction suggestion and the communicative strategies triggered by the automated feedback. These considerations will also be crossed with the observation of the texts proficiency level fluctuations during the revisions and the interaction of the subjects with the teacher through the human feedback enabled by the application. In the dynamics involving all these factors, we try to understand what this relationship implied in the writing learning process under the prism of the human-machine relationship that permeates this work.

A permanent feature in automatic writing evaluators is the focus on rewriting. *Write and Improve* is not different. By having a quick check of the text with automated feedback, the program allows several adjustments to be made in a matter of seconds, streamlining the process and keeping learners always abreast of changes and inadequacies in a visually clear and direct way. This approach to writing practice meets the pedagogical proposal on which the program is based. With the rewriting practice, the relevance in the aspect of the process stands out to that of the product, since the design of the program focuses on giving support to the improvement of the quality of writing. There is no end product; there is a continuous undertaking in the development of this skill, for there is no concern in reaching a specific number of completed texts as a finished and unchangeable work. The idea is that the learners have an environment for textual production to take place in a functional and dynamic way at their disposal, in which the interactive aspects of the software foster learning in an autonomous way,

and learners can resume their compositions at any time. In +*Class View*, the third element of the triad is the teacher/tutor/instructor, who can accompany his/her students and help them build proficiency in writing skills, complementing aspects that automated feedback does not yet provide. This triad is designed both to support the learner and to equip teacher/tutor/instructor(s) with a tool that improves the organization and visualization of the development of their students with the benefit of a prior textual revision. In the next section, we present the analysis of the interaction of each subject with the program and with the teacher-researcher.

6.4.1. Daniel

The work with one of the research subjects, who had already completed the English course at the beginning of the experiment, brought a relevant reflection regarding the use of the tool in an autonomous way. Before beginning the writing practice, there was an initial conversation to establish which topics would be relevant to the subject for textual production. After clarifying that his desire was to apply for an American university the following year and that he would have to send application letters, it was agreed that the topics proposed by the universities would be inserted in the program to become composition topics. There was no order or frequency for the productions. After registering the subject in the program, every time he produced a text, there would be feedback produced by the teacher-researcher in addition to the automated feedback from the system itself. By analyzing the feedback from the program, with suggestions and corrections, and the subject own revisions, we realized that little was understood of what the automated feedback offered. In addition, when writing feedback to the participant, we realized that the subject only made adjustments to the details explicitly addressed in human feedback. When asked about such conduct, the participant stated that he did not understand some tips and that he relied more on the feedback produced by the teacher. According to him, the information was clearer than the one the software provided. Thus, in written communication to the participant, through the *Messenger* chat application, we clarified the goals and benefits of taking advantage of the automated feedback and not ignoring it. Below there are excerpts from the teacher-researcher conversation with Daniel (Figures 35 and 36):

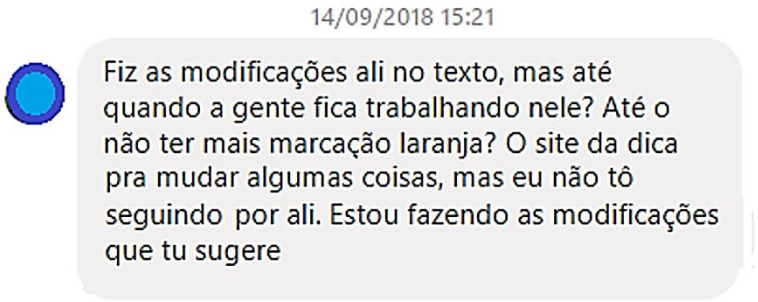


Figure 35 - Message excerpt from Daniel to the teacher regarding *Write and Improve*'s feedback on September 14th, 2018[16]

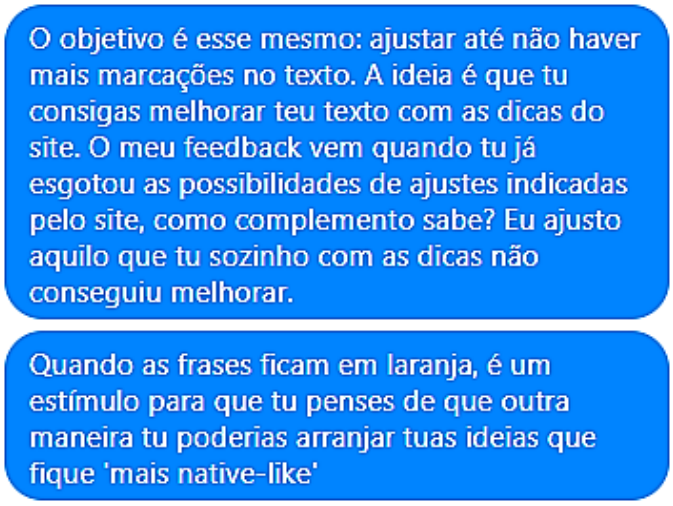


Figure 36 - Message excerpt from the teacher to Daniel about his doubts on September 15th, 2018[17]

By becoming clear to the participant how he could best benefit from the automated feedback, the adjustments to the texts were first focused on the automated tips. Only in a second moment, when his possibilities of corrections were exhausted, the participant waited for the feedback sent by the teacher-researcher, after communicating her the end of the task via text message.

[16] Our translation: *"I've made the modifications in the text, but how long do we keep working on it? Until there are not any orange markings? The tool gives tips to change some things, but I am not following them. I am making the modifications you suggest."*

[17] Our translation: *"This is the objective: fix the text until there aren't any markings. The idea is that you improve your text with the tips from the tool. My feedback comes when you have already exhausted all the possibilities of corrections pointed by the tool, as a complement, you know? I adjust those problems that you did not manage to do by yourself following the tips. When the sentences get an orange marking, it is a stimulus for you to think about how you could rearrange them so they look more "native-like"."*

Given the unprecedented use of the program by both subject and researcher, the dynamics of the experiment with the program regarding feedback provision was improved as the student reported his experiences via *Messenger* and sporadic face-to-face meetings. Thus, the format of the feedback that the teacher sent was transformed during the period, aiming to interfere less in the texts rewritings. The purpose was that the student managed to adjust most of his productions using only the automated feedback, reaching higher scores with each new version of his compositions. Next, feedback excerpts sent by the teacher-researcher at the beginning of the process (Figure 37):

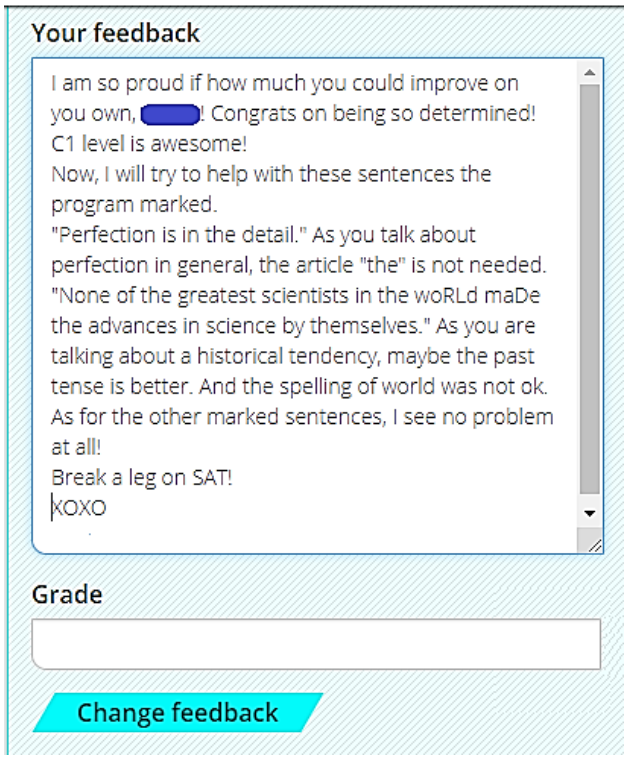


Figure 37 - Teacher feedback to Daniel on October 3rd, 2018

We should observe that the teacher-researcher transcribed parts of the text and highlighted them with quotation marks to make specific comments on each section that she felt necessary. In addition, there are words written with capital letters in order to draw Daniel's attention to spelling problems. This strategy was established in a face-to-face meeting in order to highlight this type of error since the text box for writing the feedback by the teacher/tutor/instructor does not have any type of text highlight

assistance. These problems mentioned by the researcher did not have word level feedback, only on the sentence level. Thus, there were orange markings on certain sentences that did not contain additional information. Over time, the teacher-researcher perceived Daniel's capacity in understanding metalinguistic feedback and made use of this practice to clarify some of the problems she encountered in his texts. Her goal was for Daniel to be able to remember these comments when needed by looking at his own texts and making adjustments on his own. Therefore, after copying and correcting the problematic excerpt, there was an explanation of the changes made. The following illustration is an example of metalinguistic feedback written by the teacher to one of Daniel's texts (Figure 38).

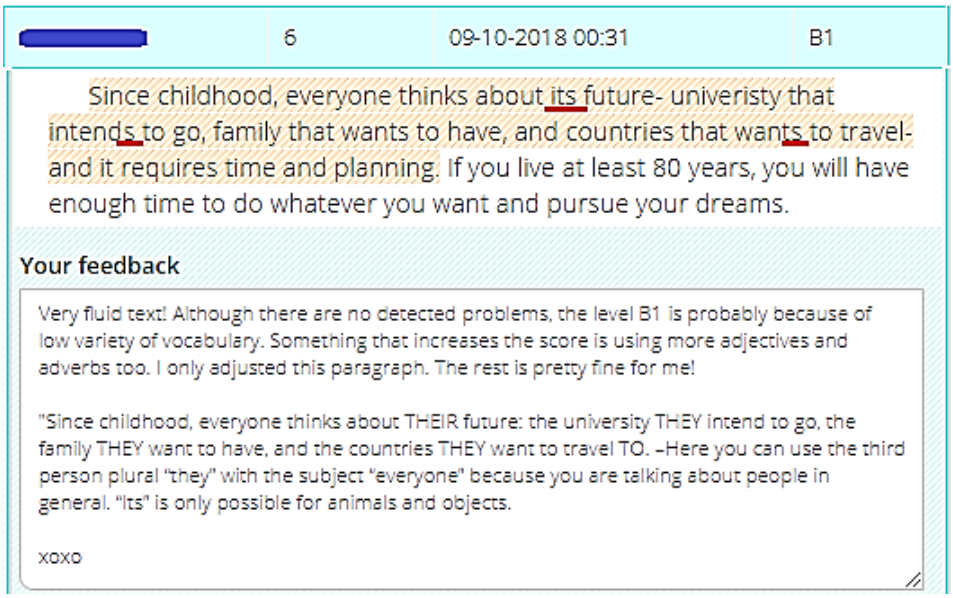


Figure 38 - Metalinguistic feedback sent to Daniel on October 9th, 2018

Another common concern of participant Daniel was about the aspects that determined the level of his texts. There was always the motivation, at the end of the adjustments in a text, to reach level C1, which he considered appropriate for his purpose of studying abroad. Therefore, reassuring and motivating feedback was needed so that the scores attributed to his texts did not frustrate him to the point of giving up practicing. As we can see in figure 38 above, the teacher begins the feedback with a compliment to Daniel's text and then tries to reassure him by clarifying possible aspects that contributed to the B1 score and what he could do differently in future versions. At the end of the feedback, the teacher brings an excerpt in quotation marks

from his text along with a piece of feedback. The teacher talks about the adequate pronoun to replace the noun *everyone*, once the learner used the possessive pronoun *its* and inflected the verbs in the third person singular, omitting the subjects of the subsequent sentences – *"[...] everyone thinks about its future – the university that intends to go, the family that wants to have, and countries that wants to travel*."

Considering that the clarification about the program's functioning happened at the beginning of the experiment, the level of his texts grew with little influence from the teacher's feedback along the process. Daniel could interpret some tips and reorganize his constructions in search of a better score for his compositions. However, the frustration was apparent when, despite several attempts to adjust the text, the level remained low or decreased. See a message from Daniel to the teacher on the subject (Figure 39):

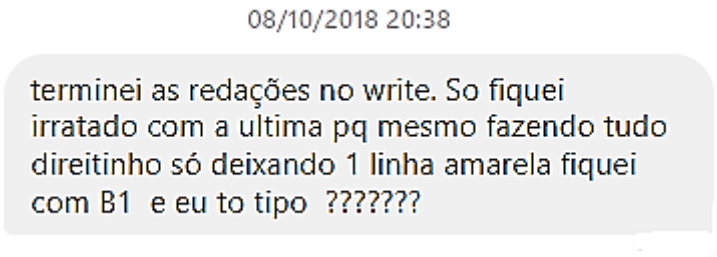


Figure 39 - Daniel's message to the teacher about one of his texts score[18]

In his message, Daniel demonstrates to understand how the feedback works and is upset with certain injustice regarding the low score assignment in view of the few markings still pending in his composition. The following is a screenshot of Daniel's original text and the corresponding feedback provided by *Write and Improve* (Figure 40):

[18] Our translation. The original: *"I've finished the compositions on Write. I've just got upset with the last one because even doing everything right just leaving one sentence marked in orange I've got a B1 and I'm like ?????"*

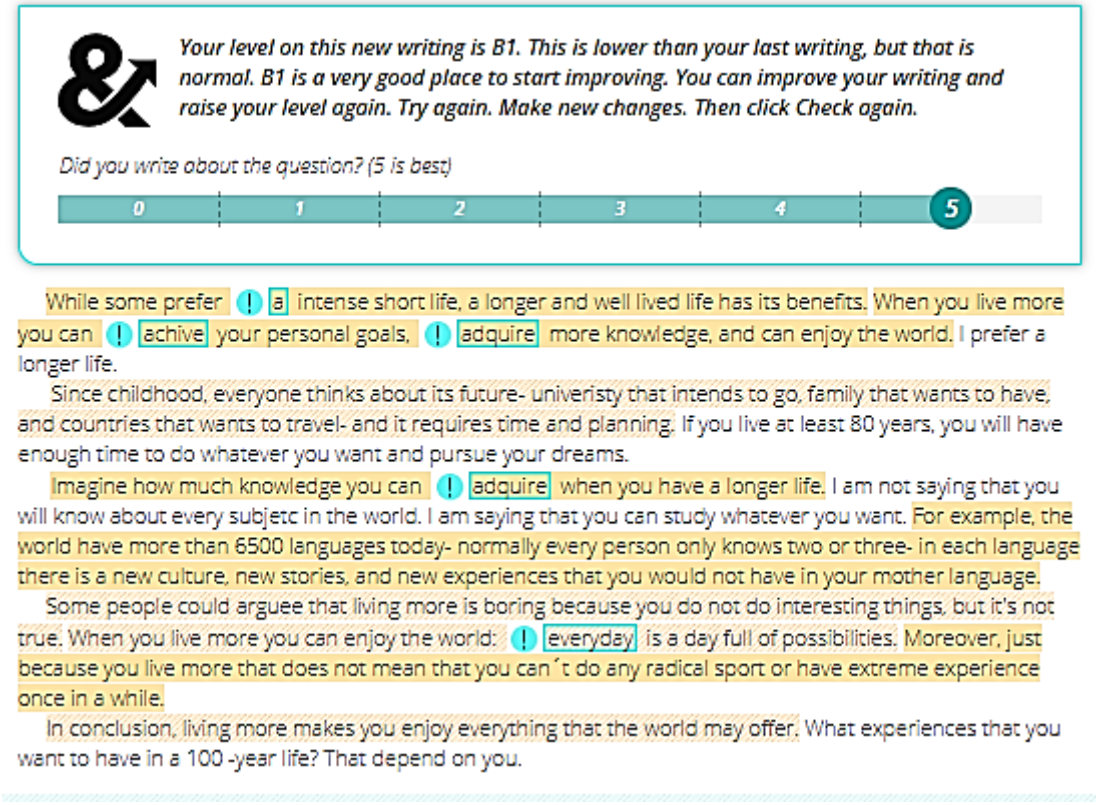


Figure 40 - Daniel's original text on October 8th, 2018

Daniel made 8 other submissions of this same text, correcting the problems pointed out and, after receiving the B1 score, he sent the teacher the message shown in figure 39. It is worth mentioning that the adjustments were made on the same day but at different times. Three of them were made in the morning and three at the end of the night. This monitoring is possible because the application records the time at which the checks are performed. This pause between the corrections indicates that there was a gap between the revisions, which suggests Daniel's concern to revise his text with some space of time in order to better observe his constructions and have a clearer judgment after a period without reading it. The teacher feedback happened only at version 6 visible in Figure 38, and Daniel revised this text 2 other times, unfortunately, without the success of a high score he had long expected. Next, the sixth version of Daniel's text (Figure 41):

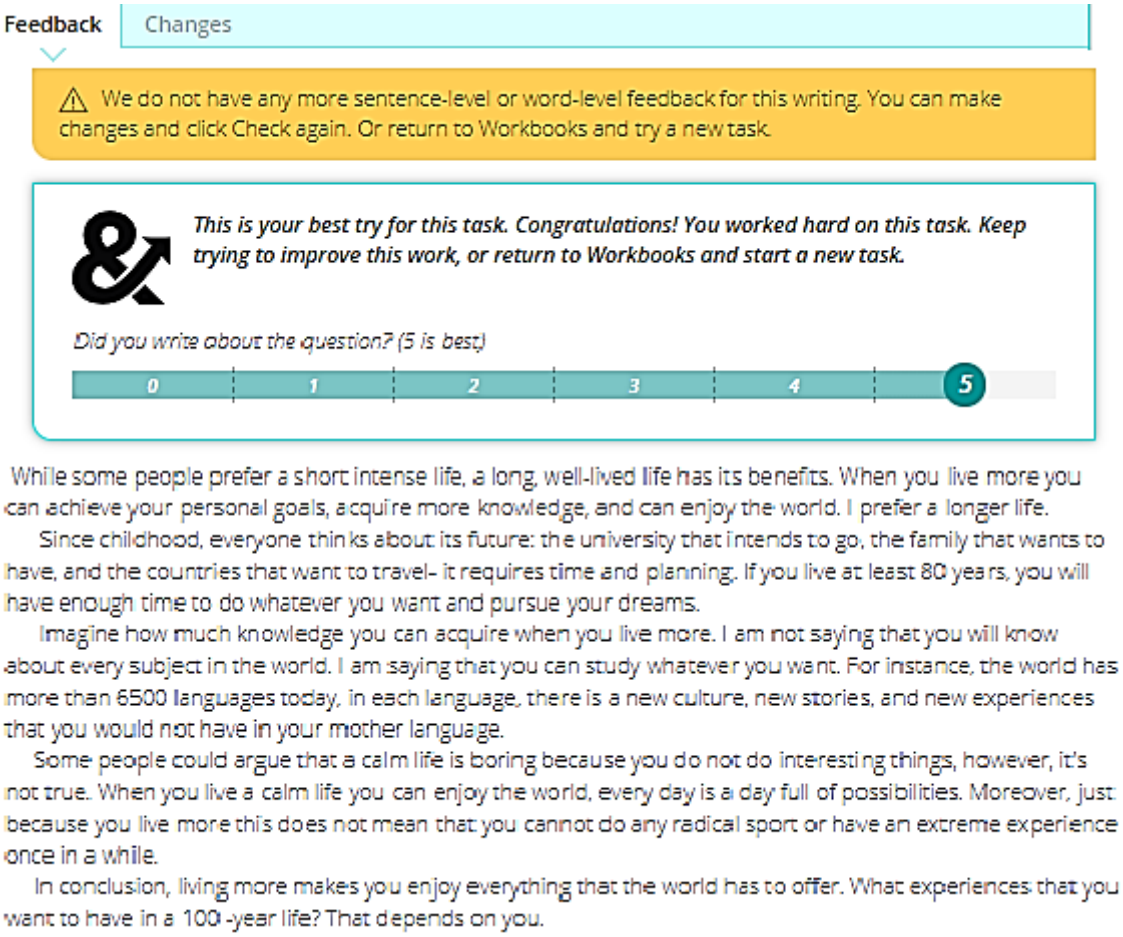


Figure 41 - Daniel's 6th version of the text on October 9th, 2018

Note that there are no markings in the text version 6. However, the level and it was assigned remained B1, which was the level of the first version. The theme remained the same, indicating that the text was kept within the proposed theme from start to finish. The overview feedback, in which motivational messages are sent to the user, brings a comment saying that this version is the best of all, which, according to Mason and Bruning (2001) is characterized by verification or KR, indicating only the text proficiency level. This happened because there is no problem in the text detectable by the algorithm, which can also be observed by the orange message at the top of figure 41, which says that there is no more feedback available on both sentence level and word level. In view of the above, this version of Daniel's text could be considered the last, since there is no apparent problem. However, Daniel is an English learner

capable of achieving C1 proficiency level in writing with support, a level already indicated by *Write and Improve* in several other compositions he wrote. For this reason, the participant was not satisfied and continued to modify his text after asking the teacher for help. Next, an illustration of the 9th and final version of the text (Figure 42):

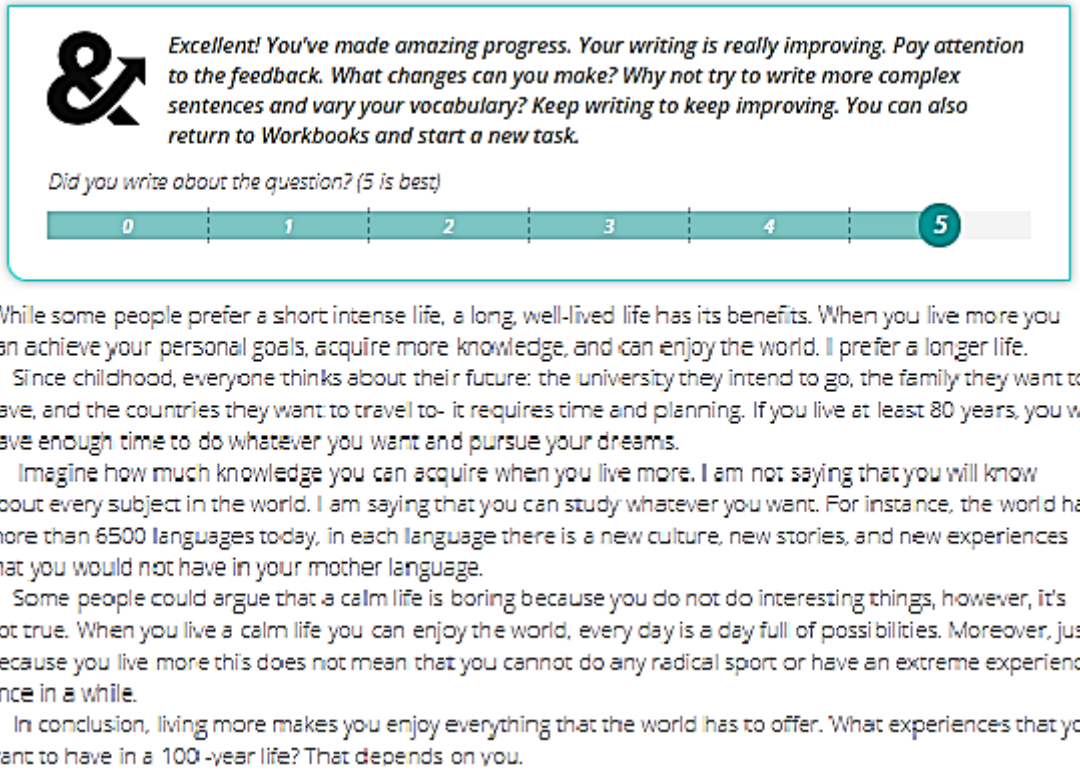


Figure 42 - Daniel's 9th version of the text on October 12th, 2018

Note that the overview feedback for this version has a very motivating message at the top, commenting on the progress made by the learner and offering tips on how to increase proficiency such as writing more complex sentences and varying vocabulary. When the message goes beyond encouragement or level assessment, it is considered formative feedback, which, according to Mason and Bruning (2001), provides the learner with tools needed to guide the execution of the task. In this case, the tips offered by the program (Figure 41) were the same as those provided by the teacher earlier (Figure 38). *Write and Improve* could have offered this formative feedback three versions back, since the text was already well constructed, avoiding that the learner experienced the feeling of demotivation, verified in the message to the teacher. Such dynamics may have negatively influenced his writing progress

perception since his writing dropped to B1 after reaching B2 in version 7. Such an oscillation in Daniel's proficiency level meets what Kluger and DeNisi (1996) discuss in their work, in the sense that feedback does not always elevate learner performance, especially when there is no specificity in the message. Finally, when the subject reached level B2, *Write and Improve* provided quite motivating feedback, as it is routine, however, no specific tip was provided anymore. Here is the overview feedback for the 7th version of Daniel's text (Figure 43):

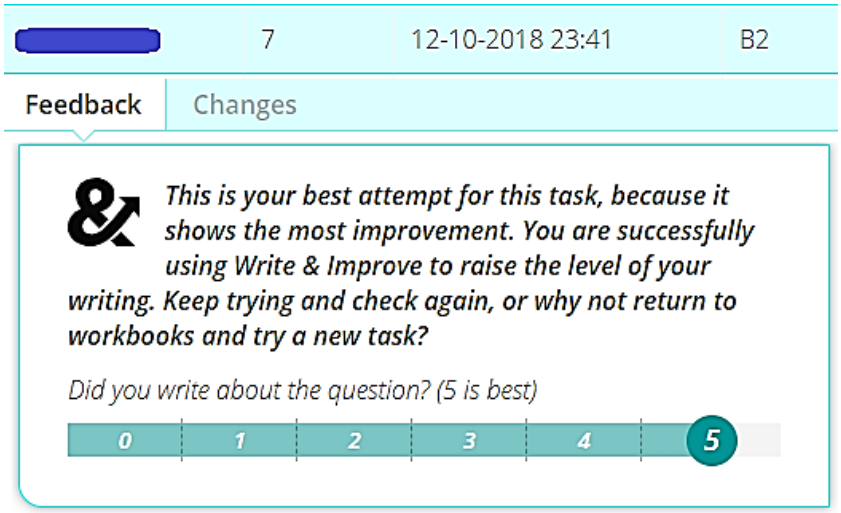


Figure 43 - Overview feedback to Daniel's 7th version of the text on October 12th, 2018

We were able to notice, with the constant observation of Daniel's texts, a proficiency level oscillation pattern. After 5 revisions with a rising level, there is a decrease again. Some variations do not move straight from one level to another because the map shows more subtle oscillations within the same proficiency level. Note that in texts 4, 5 and 6, in figure 44, the fluctuations are visible when observing that after some revisions the text level drops again, indicated by the circle of darker blue color.

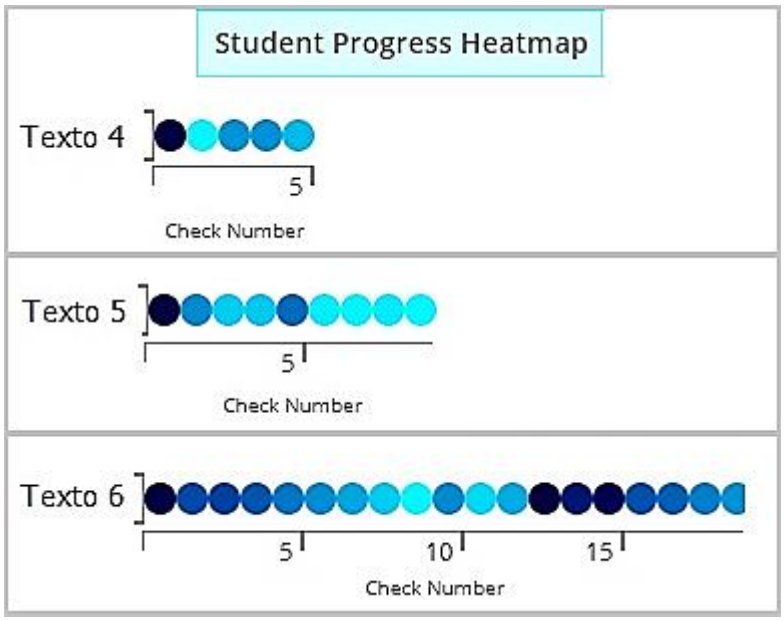


Figure 44 - Text score fluctuation from September 14th to November 8th, 2018

Score fluctuations occurred every five revisions, on average. The texts reach a higher score from their first version around revision number 5 and after around other 5 revisions there is a drop in proficiency level, and in the sequence the level fluctuates again, reaching the initial score again. One of Daniel's texts, for example, moved between levels B2 and C1 during 41 revisions, in what can be understood as a wave, with apparent peaks and valleys, but at no time surpassed level C1 even with numerous adjustments. See the chart that the program makes available for this aspect (Figure 45):

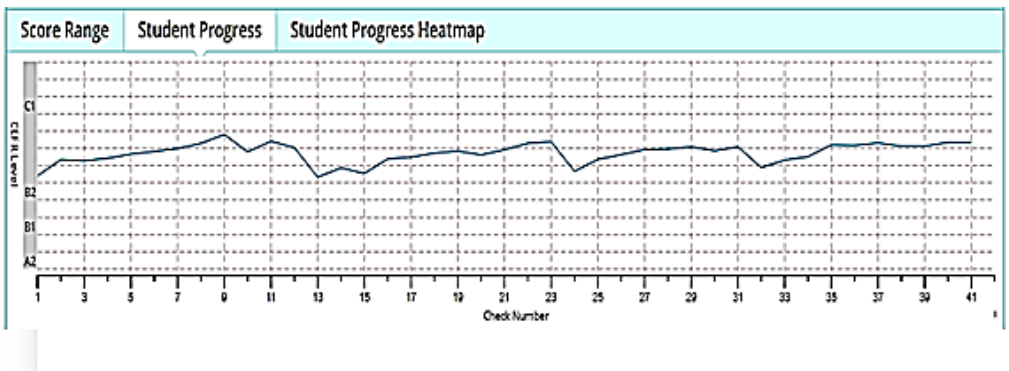


Figure 45 - Daniel's progress graph of one of his texts

Daniel's early texts always moved from level B2 to level C1. Under a constructivist view of learning, it becomes evident that B2 is his current language proficiency level. However, with the automated tool support, his assisted performance, which is encouraged by *Write and Improve*'s feedback, reaches level C1. There is an indication of his ZPD. From this angle, *Write and Improve* demonstrates effective support in the development of the learner's writing skill, leading him to reach his maximum potential of language proficiency.

Looking at the rewriting perspective, the waveform line on the proficiency graph shows that the learner's text does not improve due to excessive revisions, because in 5 or 41 revisions, Daniel reached level C1 and did not exceed it. We can understand from this outcome that it is not the time devoted to a particular text that will necessarily entail a more proficient writing level; it is the quality of that time in the sense of the strategies undertaken in solving linguistic problems with appropriate assistance. This oscillation in written language proficiency further shows that learning, in fact, is not linear and upward. Daniel's interlanguage behaved heterogeneously even with the immediate action of automatic feedback in his texts over time.

In another text of Daniel's authorship, the interaction with the program was curious when we observed that *Write and Improve* provided feedback with little coherent suggestions considering the syntactic context of the sentence where the problem was located. However, although the automatic tip was irrelevant, the communication was established visually, with the color marking that the system offers. The following is an illustration of the program marking and feedback for the error (Figure 46):

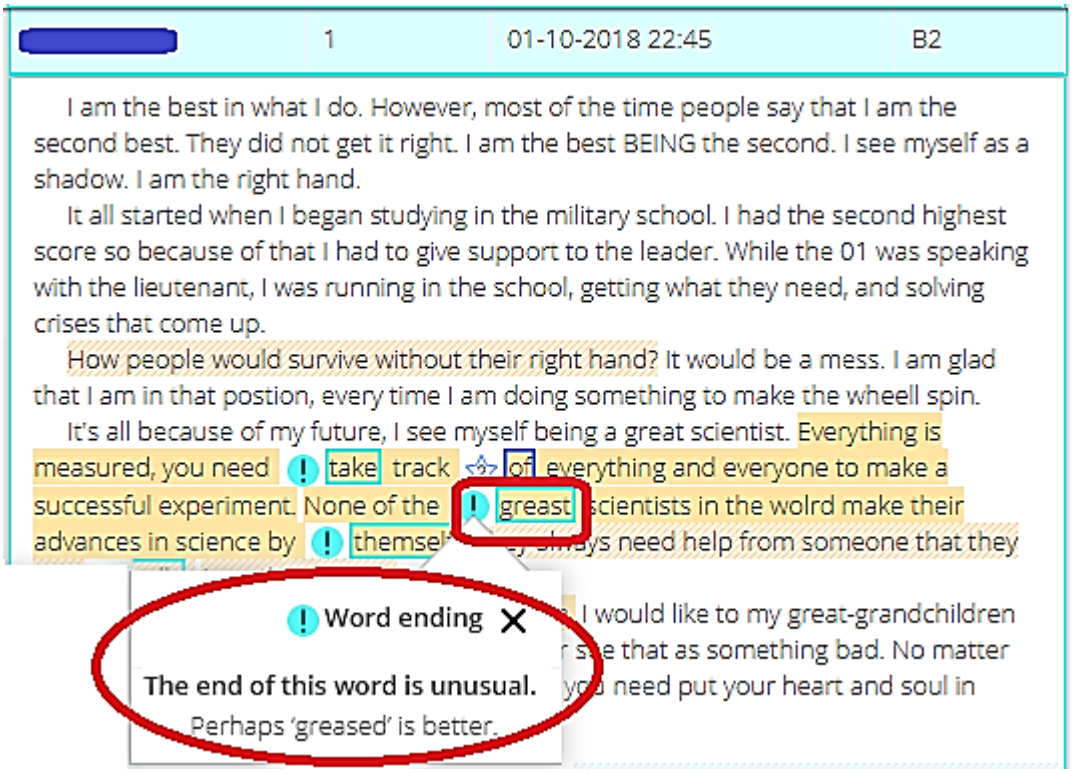


Figure 46 - Problem marking and automated feedback in Daniel's text on October 1st, 2018

We should mention that the automated hint suggests that Daniel replace the word *greast* with incorrect spelling for the adjective *greased*, which little approaches with any idea that he developed within the sentence. Daniel tried to produce a construction in the superlative and had problems, possibly a typo. The program located the wrong word in isolation from its syntactic context and suggested the word considered closest to that spelling. We can also observe that the feedback draws attention to the end of the word (*word ending*). If we take into account the question of words ending in *ed* in English, there is a class of them that are pronounced with voiceless sound and others with voiced sound, especially in the case of verbs inflected in the past tense, for example. However, *Write and Improve* does not work with the phonetic aspect of the words, which would make this association not a possible way to understand the origin of the suggestion.

On the other hand, when typing in on the Cambridge dictionary online the construction *greast*, because it is not in a fact word, word options appear as approximate suggestions. The dictionary suggests 10 similar words and *greased* is in 8th place. The word *greatest* appears as the 5th option, which also does not elucidate

the origin of such a tip. The fact that the word was corrected to *greatest* in the next version indicates that Daniel took advantage of the feedback, realizing that the marked word was not written as he intended and discarded the feedback as being irrelevant to the problem.

In this same text, another similar feedback occurred, where the suggestion of correction was far from the text topic. The participant possibly had another typo, writing the word *relly*, having intended to produce *rely*. *Write and Improve* suggests a correction – *really* –, which would turn the sentence into an ungrammatical construction. Again, the word was captured in isolation and little helped the learner. When the researcher asked Daniel about this feedback, he replied that he realized by means of the blue marking he had a typo and searched for the correct spelling of the word he wanted to write. As for the application suggestion, the learner simply discarded it.

Given these two examples of word-level feedback that *Write and Improve* provided, we can notice that some suggestions still do not foster improvement in the writing of the text, serving only as markings of irregularities that learners adjusts on their own. The algorithm model offers unsatisfactory hints due to the way it is trained to find approximate words to those written incorrectly in the text. In summary, for each unidentified construction, the algorithm encodes it as OOV (*out of vocabulary word*), and from there, searches the origin of the construction through an unsupervised word aligner. This means that the algorithm looks for the root of the word and tries to suggest other approximate words registered as possibilities in its database (BRISCOE; YOUN, 2016). As we can see, this choice process is fully mechanized, that is, there is no human judgment in the selection of the suggested word. Thus, suggestions do not always make sense when looking at sentences as a whole. In this matter, taking this aspect of the system into a more philosophical ground, Weizenbaum (1976) states that judgment is what makes us uniquely humans and at the moment machines get to be in charge of that we put at stake the essence of what it means to be human. Therefore, we should not expect the system to execute such tasks sensibly.

Another relevant point is to remember that the program was not trained to mark all the errors present in a text. This is because developers believe that very marked text has a negative effect on learners, who may feel frustrated by seeing few of their constructions free of markings. However, even if word-level and sentence-level feedback ceases to appear, the final score of their text will indicate that not everything

is appropriate. This fact was evidenced by observing Daniel rewriting his texts several times and even though there were no markings, his score was not what he expected. When reading the texts carefully, the teacher encountered problems that could have interfered in the assessment. By correcting them and resubmitting the text for feedback on her own profile, the text reached a higher score. So, even if there is no marking yet, the work with the application can become exhaustive for learners, as they will find themselves without pedagogical guidance to reach their maximum potential. There is an indication of the end of productive interaction with the program, where it reached the threshold of providing automated assistance to users. From that moment on, learners, in the desire to continue perfecting their text, will have to seek other resources and supply their unmet needs.

On the other hand, if we consider a very marked text somewhat frustrating, and this is a concern on the part of the application developers, the practice has shown that an unmarked text with the still below-expected score also causes frustration since learners can drift without guidance on what to improve. Working with *Write and Improve* proves to be a challenging task, in the sense that we can miss rich learning moments by not rewriting the text any longer trying to perfect it. That is the precise moment when teacher involvement is an essential aspect, bringing students' confidence back, offering linguistic as well as psychological support to deal with the disappointments and the challenges that follow the process of learning a new language.

To deal with the problems pointed out by the system, a strategy largely used by Daniel was substitution, deployed to react to word-level feedback. One of the examples of sustitution was the collocation *good at* (Figure 47), adjusted by trial and error since he tried two different prepositions until the marking disappeared. This strategy was needed because of the program's vague feedback, which pointed an inadequacy by saying that the word did not seem right.

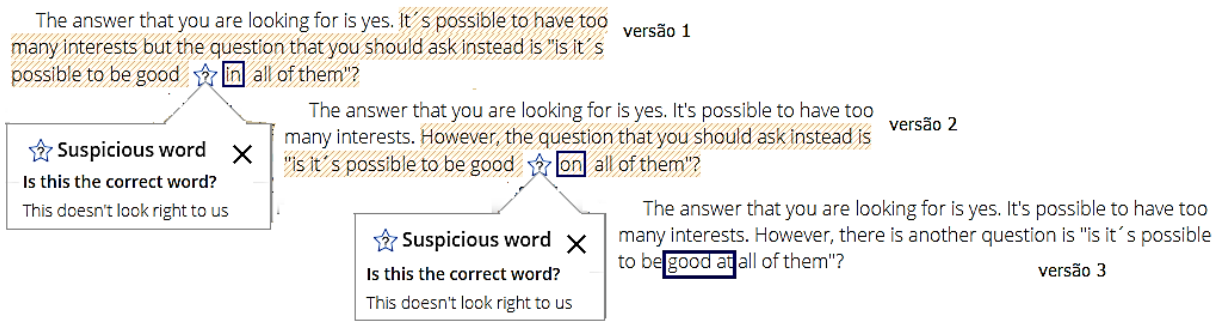


Figure 47 - Word-level feedback to the collocation *good at*

In version 1, Daniel used the expression *good in*, possibly influenced by his ML, demonstrating that Daniel made a literal translation for that first attempt. Not being successful and receiving vague feedback from the system, the learner restructures his expression by changing the preposition *in* to the preposition *on* in version 2. Again unsuccessfully, Daniel changes the expression one more time, switching *on* to *at*, visible in version 3, at which time the program stops marking the text, informing the learner that the preposition *at* is the most appropriate.

Figure 48, below, also shows another very recurring problem among Brazilian variant English learners using the verb *take* for the expression *take a decision*. This expression accompanied by the verb *take* is not incorrect, but it is not the most common. In the search for the volume of use of the two constructions on Google search engine, the expression composed by the verb *make* gets a higher number of entries. Native speakers in online language forums consider the meaning of the expression with the verb *make* indicative of pondering the situation on which to decide something, whereas "*take a decision"* relates to the act of implementing the decision made. This is considered a probable interference of Daniel's ML, in which he uses literal translation strategy to communicate his idea of "*tomar uma decisão*", since the verb *take* can be translated as "*tomar"* Portuguese, among others. The use of the verb *take* was pointed out by the program, but there is no feedback offering suggestions because of this double possibility of construction. The program did not go beyond marking the verb of the expression as suspicious, leaving it to the learner to find out what the best option for his sentence would be, thus avoiding an inappropriate correction.

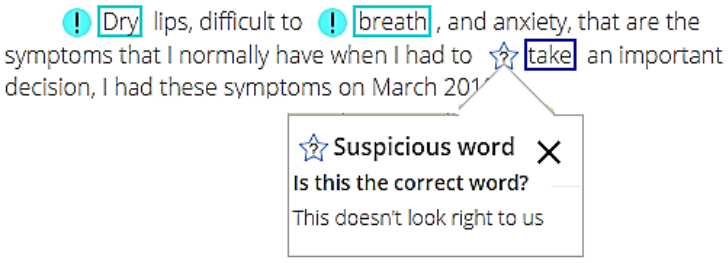


Figure 48 - Feedback to the inappropriate use of the verb *take* on September 11th, 2018

For errors involving the ML transfer process, there would be a need to train the algorithm to identify problems arising from this aspect. However, such an approach to providing automatic feedback requires the development of applications with a level of complexity yet to be achieved, especially because of the need for extremely detailed data classification. While this is not a palpable reality, we follow the data analysis in which the feedback was favorable to the learners, offering suggestions consistent with their needs.

Another type of common language inadequacy among learners refers to the nominal agreement, where pluralization of words is not well-constructed throughout the sentence. When the algorithm identifies this problem, the feedback provided is called *agreement*, presented in a clear format, making it easier for the learner not only to correct the problems but also to understand the source of the inadequacy. See figure 49, where this type of feedback appears in one of Daniel's texts and note the complexity of marking such a problem:

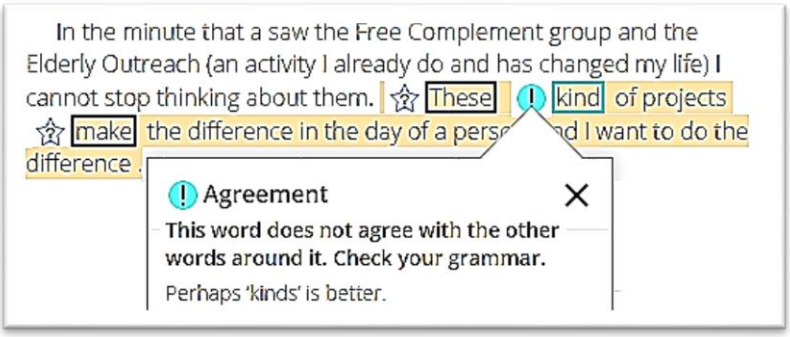


Figure 49 - Feedback to one of Daniel's text on October 1st, 2018

The identification of a problem of this nature occurs with the algorithm marking all the words that need to be modified. However, they are presented with distinct markings. In the sentence "*These kind of projects make the difference in the day of a person and I want to do the difference.*", the program marks the determinant *these* as a suspicious word. In the next word, marking is at the core noun of the subject, with feedback on agreement problem. The feedback message explains that the word does not agree with the other words around and suggests its pluralization – *kinds*. In the sequence, there is an indication of a problem in the verb construction and the feedback presented is *suspicious word*, because if the learner kept the sentence with the subject in the singular form, the verb would have to be otherwise inflected, becoming *makes*.

We can verify in this example that the algorithm looked for the most practical solution to the problem since the only word that did not agree with the construction was *kind*. However, the other markings, though pertinent, become clearer to someone with more extensive grammar knowledge. Thus, it seems more sensible that learners focus first on the more direct and clear feedback and leave vague feedback for later revisions. This observation is possible because, when looking at Daniel's rewritings, this was his strategy: he corrected the term *kind* according to feedback, and in the following version, the other markings were no longer present. Therefore, a single adjustment eliminated three other markings in his sentence, in addition to the orange marking. It is worth noting that the participant had not received any kind of instruction from the teacher on how to prioritize the adjustments suggested by the automatic feedback; only the guidance that he should seek to eliminate all the markings present in his text to achieve a good level in the program's assessment.

Still examining this same sentence written by Daniel, we perceive an inadequacy of language that in no moment received feedback. It is the final part of his construction in which there is *"[...] I want to do the difference."* Here, the verb used is not appropriate for the idea intended. The proper verb would be *make*. As the algorithm did not address this problem in any of the 41 times Daniel rewrote his text, we believed that such knowledge would not be acquired. Daniel was not alerted, not even by indirect feedback, that there was an inadequacy in this part of his construction, which does not represent natural use of the language, even when there were no other problems present in his text. However, the first part of his sentence had received feedback on agreement that eventually marked the verb *make*, which is part of the initial construction "*These kind of projects make the difference [...]*". We understand, therefore, that there is the possibility that even when there is no direct feedback for a given problem, the markings work as a visual alert, drawing the learner's attention to another inadequate form, as we believe to have been the case in this example.

Since Daniel had made the correct use of the verb at first and then produced the same expression in an erroneous way, we notice the presence of a slip and not of a form not yet acquired. Having said that, it is necessary to point out that in terms of accompanying the process of writing development, *Write and Improve* offers the possibility for the teacher/tutor/instructor to visualize the stages through which the learners go and the strategies that they apply to overcome the difficulties that their interlanguage, mental linguistic system still in development, imposes on them. In

addition, the organization of the records of each rewriting submitted to the program assessment can facilitate the association of certain linguistic behavior with other factors, extralinguistic or not, present in the learning scenario of the subject and that can influence the process.

In another excerpt of Daniel's text, there was another rewriting process that illustrates the triad of interaction among the learner, the program, and the teacher, in which the subject works with all the input he receives to improve his composition. The final part of Daniel's text, which brings his opinion about a particular university, underwent several modifications until it was free of markings, which may indicate different learner behaviors in relation to the feedback he received. The original passage is "*Everything what Tufts represents is the reason that makes me smile every time when someone asks* "*what college are you applying?*" and the final version is presented as *"When someone asks me what college I am applying, I always smile My answer is Tufts*."(Figure 50).

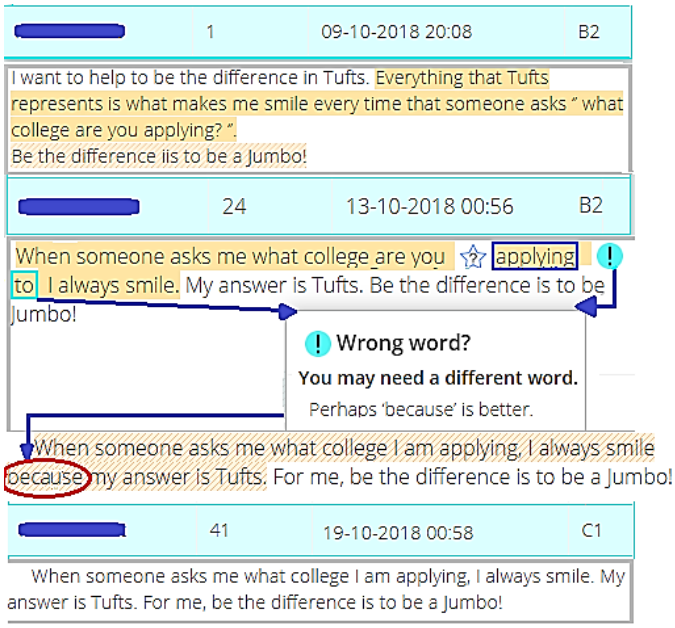


Figure 50 - First, intermediate and final version of Daniel's text excerpt on October 9th, 2018

At first glance, we can believe that the learner realized he was using the indirect discourse in an inadequate way besides splitting his extensive subordinate clause into two other sentences. However, it took 28 revisions to arrive at the final version. On the way, there was an intermediate version *"When someone asks "what college are you applying?", I always smile because my answer is Tufts.",* composed of two sentences

with the addition of *because* in the final part. The indirect discourse and the dissolution of the sentence were only corrected later with the support of the teacher's feedback, since the algorithm did not provide a clear feedback for this problem, only marking the construction in orange. Even the teacher having rewritten the excerpt to the subject, his next version was not completely adjusted, being in this format: *"When someone asks me what college you are applying to I always smile. My answer is Tufts*." There was only one interaction of the teacher with the learner via the application, and the other corrections were entirely made by the subject himself, trying to reformulate his constructions and raise his text score. The score varied between B2 and C1 during this process and this excerpt received automatic feedback due to the preposition added after the verb *applying*, suggesting the change to the conjunction *because*. Daniel did not comply to the suggestion since his following version presented only the omission of the preposition: "*When someone asks me what college I am applying I always smile. My answer is Tufts.* " Still, the program marked the sentence in orange, pointing out that improvements could still be made. After 4 revisions, the learner decided to add *because* in his construction in an attempt to eliminate the orange marking, but without success. Then, Daniel removed the conjunction and added a comma after the term smile and submitted the text again, and this time he got a sentence free of markings.

Automated feedback should foster interaction so that the communication between the program and the learner happens. This aspect is crucial to trigger learning. Without this factor, the interaction is not established, the communication fails and the goal of the pedagogical tool is lost. The question that revolves around this process is how to consolidate it when designing a digital product for language learning. It is not an easy task and there are no solid standards for evaluating which tool design is most effective. It is a set of factors that must be analyzed simultaneously. We must consider aspects such as who will use this system, the pace at which the interaction will happen, what results are expected as well as who will be in charge of coaching this pedagogical process.

Autonomy is undoubtedly a path being sought in the 21st century. As a result of the new cultural objects inserted in human daily life that organize the course of learning in an automated way, it is visible that they interfere and over the decades cause transformations in the acquisition of knowledge and consequently in FL development. It still seems too early to establish a new paradigm and its possible implications. However, it is important to follow the movement that technology-mediated teaching-

learning is undertaking, associated with the use of artificial intelligence in the search for the optimization of pedagogical work in a posthuman perspective.

6.4.2. Beginner level participants

From the 3 subjects that participated in this research, the 2 girls were beginner-level learners. Also, they were part of one of the teacher‘s class at the private course she worked. This way, she could maintain regular contact with the students, following more closely their interaction with the program and this facilitated the adjustment of obstacles found during the process. Therefore, there was the possibility of having a workshop with the class, in the computer lab part of the dependencies of the course where they studied. The workshop lasted two hours and each student had a desktop computer available to perform the tasks with the support of the teacher. The workshop was designed so that they would have a moment to practice the writing skill in the application with the help of the teacher in a synchronous way, solving doubts and discovering how to better interact and take advantage of the automated feedback that *Write and Improve* provides.

On the other hand, the teacher also benefited from the event once she was able to observe closely the effectiveness in the use of the application, diagnosing the gaps between what the program proposes to offer and what the students desire to find when performing this type of activity. Likewise, it was possible to evaluate the functionality of the feedback in establishing communication with the learners and giving the assistance they needed according to the level of language they had, measuring how much and in what aspects the teacher would need to interfere so that the flaws in interaction were attenuated.

6.4.2.1. Jaqueline

The participant Jaqueline performed 6 of the 10 tasks proposed, having a total of 24 productions, a number that includes all the versions of her texts. From the 9 tasks, 4 were done before the writing workshop, and the other 5 tasks were completed on the day of the workshop. After the event, the participant did not use the application anymore, even with classes still taking place, possibly a result from the low interactivity

offered by the application to beginner learners once it is fully configured in the target language.

Her first composition was entitled *Emergency situations*, having only one rewriting. In this corrected version (Figure 51), we observe adjustments to direct feedback, which indicated spelling problems. The second composition was entitled *Describing Your Hometown* and, in addition to spelling adjustments, the participant also used a strategy to solve a structural problem indicated by the automated feedback. See below for the inadequate construction still uncorrected:

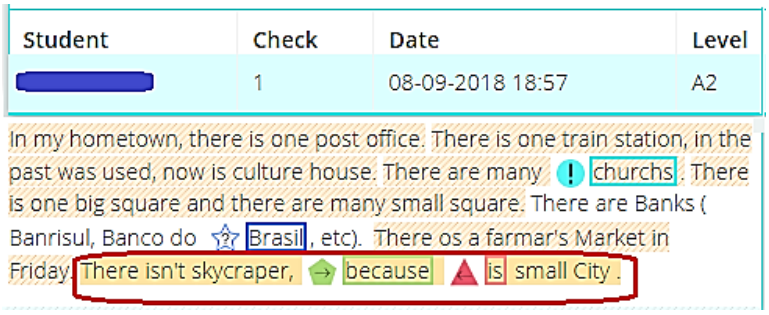


Figure 51 - Jaqueline's text on September 8th, 2018

Jaqueline received feedback for the construction *"There isn't skycraper, because is small city."* The pentagon shaped symbol in green containing an arrow pointing to the right alerts for the lack of a word after the word marked with a green outlining – the conjunction *because*. Next, there is a triangle shaped symbol in red containing an arrow pointing to the left, which refers to the lack of a word before the word outlined in red, in this case, the verb *is*. Therefore, we can see that the two markings refer to the same problem – the lack of a word to fill the subject position of a coordinate sentence. In addition, there were message boxes indicating the problem, and the suggestion was to insert the pronoun *it* in both markings.

The solution found by the learner was to substitute the verb *is* by the pronoun *it*, probably influenced by the red color marking that the verb obtained. We understand by means of this case that the learner had a misinterpretation of the visual presentation of the feedback, which led her to reconstruct an inappropriate sentence, even though she made the suggested correction. The restructuring strategy that the participant undertook was directly affected by the limited interaction and consequent failure to communicate with the program, compromising the improvement of the text.

We cannot say how the reformulation of this sentence would have been if it had been carried out without the application mediation. Learning is non-linear and unpredictable, being able to follow different interaction paths with several internal and external elements. However, this is a completely new scenario in which an automated system interacts with the human learner to enhance the construction of a text. That being said, there is still no taxonomy in the literature of the area that exploits communicative strategies under these conditions since the restructuring that the participant has resorted to fits only partially in the category of restructuring, which does not presuppose the direct visual and linguistic interference that the automatic feedback produces.

Words with incorrect spelling are the main mistake made in the program due to the typing process and consequent inexistent proofreading before submitting to feedback. When words with incorrect spelling are positioned as the subject of a sentence, the problem is not identified by the automated system as easily as in other positions. Jaqueline, for example, produced *Shoud* in one of her texts, with the intention of producing *Should*, and the algorithm was not able to make satisfactory correlations and provide feedback that could help her perceive this incoherence. As a result, the program identified *Shoud* as a proper noun and did not mark it. The complete construction was "*Shoud speaking for me and your familly your preocupation*." We can notice that there were other problems in her sentence, however, only the spelling of the words *familly* and *preocupation* were identified, marked and provided direct feedback, that is, suggesting the correct spelling. After adjusting these words, the sentence remained with the orange shade and Jaqueline had no other tangible tips on how to proceed with her production. Note the illustration of the learner's text and the respective markings (Figure 52):

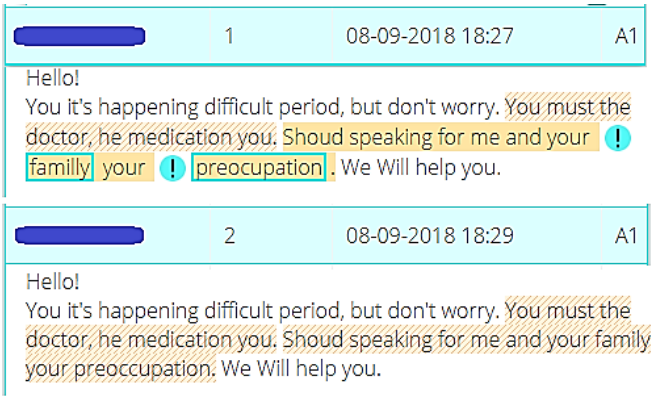


Figure 52 - Jaqueline's text before and after suggested corrections on September 8th, 2018

Because of this reaction of the application, not marking the word *Shoud*, we have the evidence of the position of the subject in the sentences receive differentiated care, in the sense that misspelled words can be interpreted by the algorithm as proper nouns. Because *Write and Improve* is used by English learners as FL, proper nouns of various origins can arise in the subject position of sentences. Therefore, if the algorithm were to mark all the non-English names present in the compositions, there would be a considerable amount of inappropriate feedback, which would have a negative impact on learners' interaction with the program. Thus, we understand that the command of the algorithm is not to mark words directly at the beginning of sentences, even if these words are incorrectly written.

The texts presented so far were written before the writing production workshop that the teacher had with Jaqueline's class. Jaqueline's texts, then, were all written during the workshop, in which the teacher was present and offered constant assistance.

Jaqueline's first production in the workshop was a rewriting of a text she had made a month earlier, on October 21, 2018 (Figure 53). On the day of the workshop, she read the feedback written by the teacher days earlier and tried to solve the orange markings as suggested.

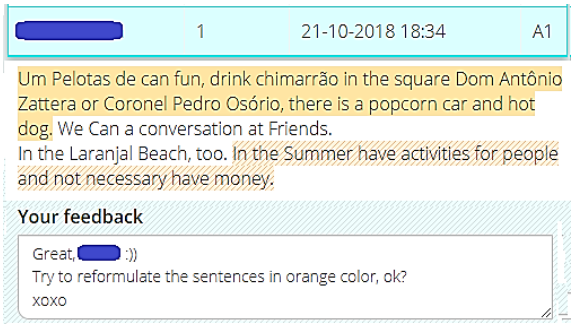


Figure 53 - Jaqueline's text on October 21st, 2018

Jaqueline was able to solve some problems even with only the indirect feedback from both the application and the teacher. In the next version of her text, the student noticed the incorrect spelling of the preposition *in*, which she typed *one*, right at the beginning of her text. She also noticed a word from Portuguese present in her first

sentence - *de* -, which she removed. Realizing that the orange color had faded, the teacher greeted her for her good performance and motivated her to try to solve other possible problems in her sentences so they would be unmarked.

However, even after 35 minutes, the student was not able to seek other possibilities for her constructions, asking the teacher for help. Beside the student, the teacher pointed to the problem of using the verb to indicate the idea of existence, in an attempt to activate the memory of the learner for a topic already worked in the classroom and on which Jaqueline had already produced a text. The attempt was successful and the participant rewrote her sentence with the structure *there is*. However, the marking remained, because the verb needed to be inflected in the plural form. The automatic feedback did not fill an important gap and if there was a metalinguistic hint, the learner would have probably been able to adjust her text without resorting to the teacher. However, in spite of the simplicity of the error, the algorithm failed to give continuous support to the participant, which places its effectiveness below the expectation from the teacher's point of view.

As for the parts of her text that included proper names, in the case names of two squares from her hometown, Jaqueline had been warned that there was a possibility that the marking of feedback might happen by not recognizing these specific nouns and that not necessarily her sentence was incorrect. The participant decided to remove the terms to make sure there were no problems other than that. However, there was no success in this endeavor, since the text continued with a marked construction. Thus, even with the help of the teacher in restructuring the sentence, the student gave up due to lack of subsidies to reformulate her sentence. As follows, the final version of her text (Figure 54).

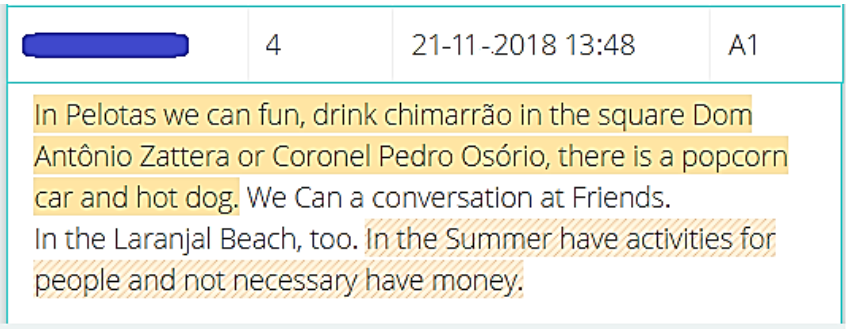


Figure 54 - Jaqueline's final version of text on November 21st, 2018

The teacher, after the workshop, worked on the student's constructions in her own application account in order to build a sentence that was not marked. Similarly, the attempt was not successful. This manifestation of the program happens because it is not as user-friendly for beginner learners as it is for learners at a more proficient level because the algorithm requires slightly longer texts to carry out a more objective and clear work. In the meantime, the application is available free at the elementary level as well, and this information is not posted on the page to alert those who are venturing for the first time in writing in English. This peculiarity is accessible to the public only through a web page linked to the frequently asked questions program, the FAQs. Users confused with feedback asked questions about their text assessment in relation to the level at which the task they performed was assigned and the answer to that question is the following:

> The tool classifies writing, not you. Although you may be a proficient English user, if you answer a question in the Novice workbooks, it will be a good standard essay for beginners - a short essay, with no complex grammar, simple sentences, but with a good range of beginner vocabulary. And an essay you write in Advanced will be the Advanced standard. It is impossible to score C1 for a beginner essay, because to score C1 you need to demonstrate a wide range of grammatical structures and functions and complex vocabulary and phraseology. You can not do this in 30 words about your daily routine, for example. It's important to choose the right workbook for you."[19]

6.4.2.2. Jessica

The participant Jessica performed all 10 tasks proposed, having 37 productions, a number that includes all the versions of her texts. From the 10 tasks, 4 were done before the writing workshop, and the other 6 tasks were completed on the day of the workshop. After the event, the participant did use the application anymore, even with the classes still going on.

Jessica's first composition was entitled *Emergency situations*, and there were no rewritings. The teacher provided written feedback via application, but the learner did not go back to her text later. During one of her classes, when asked about such behavior, the student considered it unrewarding to rewrite her texts and check the feedback the teacher provided. The tasks were performed more for the sake of fulfilling

19. Retrieved from https://help.writeandimprove.com/frequently-asked-questions/i-am-a-c1c2-speaker-but-the-essay-i-wrote-in-the-beginner-workbook-was-scored-a2-why in Nov 2, 2018.

a requirement than due to intrinsic motivation to improve her writing. Along with this statement, the low level of interactivity conferred by *Write and Improve* may have influenced the learner's lack of interest in the writing practice.

The participant performed 3 other tasks during the approximate time of one hour, one titled *Little money, much fun* and another with the title *Describe your hometown.* The last task was called "*Describe a romantic, funny or scary event in your life".* Among these tasks, the latter obtained four rewritings without any type of feedback provided by the teacher. Although her text did not present any spelling problems or other inadequacies, Jessica tried to rearrange the sentences marced in orange in order to have a more correct text. It is worth mentioning that the learner had not been instructed on the approach that should be used in the interaction with the application. Intuitively, Jessica understood that the feedback was showing linguistic problems and worked on solutions, restructuring her sentnces. The fact that the theme of the composition and the requested structure – simple and progressive past – having been worked on at a previous moment in the classroom may have influenced the learner's behavior. Here is how Jessica handled the sentences marked (Figure 55):

Figure 55 - Jéssica's text on August 6th, 2018

We must observe that no rewriting was objectively marked with the standard symbology, but because of the orange coloring, the learner was led to believe that the problem was in the verbal inflections, since these were changed every version.

In the first rewriting, that is, in the text number 2, Jessica modifies the inflection of the verb *to be* in the past, from third person singular to third person plural and removes the end of her sentence. This strategy used by the participant is restructuring, in which the learner restructures her construction by changing some elements, mainly verbs. In addition to this strategy, the student also makes use of message reduction, in which there is a removal of terms. Thus, her final sentence becomes *"Thiago were drinking beer."* Without success, the learner reformulates her sentence again, adding other elements like the conjunction *when*, creating a subordinate clause, visible in version number 3 of her text, probably imagining a necessity of this element for descriptive constructions, a topic also addressed in class. However, her sentence receives feedback with an intense orange coloration, which means a more problematic construction than the previous one, although clearly more complex.

In the sequence, in text number 4, Jessica returns to the singular third person inflection for the verb *to be*, probably realizing that this could be the problem created in version number 3. There was also the change of the main verb of the sentence, from *drinking* to *making*, a replacement that may have been motivated to avoid repetition of the verb in the same construction. This exchange of elements is called substitution strategy, in which one term is replaced by another in an attempt to obtain a totally correct sentence. It is likely that, intuitively, the student thought that it was impossible for the program to recognize words from the inventory of the Portuguese language and to acknowledge its use as an error and, to eliminate this possibility, relied on this mechanism. It is interesting to note that the noun *churrasco,* also present in the text, at no time caused doubts to the learner since the sentence that contained it, although inadequate, received no feedback and, consequently, received no attention from Jessica.

Once the sentence was diagnosed as problematic, the learner reformulated her sentence in 4 different ways in search for positive feedback, that is, a sentence free of markings. However, without success with the strategies employed, and without the

ability to use others, Jessica abandoned the text and turned to another task, even after the teacher had provided feedback through the application, rewriting her problematic sentence.

The versions and modifications that Jessica made in her text point to the use of restructuring, substitution and reduction as communicative strategies, indicating that the learner was presented with manageable units of feedback from the program, triggering the use of such strategies. We also observed that, although Jessica's texts did not fluctuate between the CEFR levels – all four versions scored A1 – *Write and Improve* offers a graph in which we can perceive more subtle oscillations in proficiency levels (Figure 56).

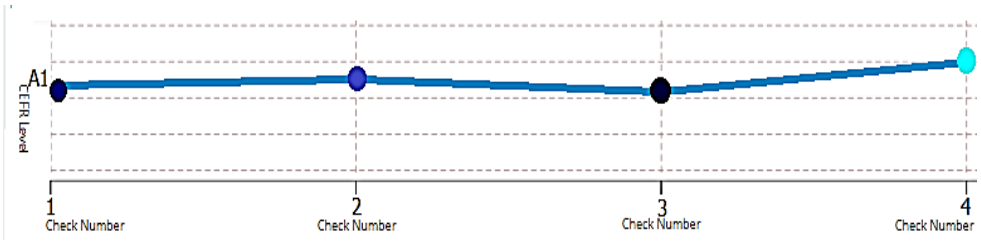


Figure 56 - Jéssica’s progress graph for one of her texts

It should be noted that the version that reached the highest level was number 4, version in which, although there are words that do not belong to the English language inventory – *churrasco* and *caipirinha* – two proper names and other unmarked inadequacies, contained a subordinate clause and all verbs present in the text were inflected correctly. However, version number 2 also achieved a high level of proficiency despite having an incorrectly inflected verb.

Comparing the construction of version 2 and version 3 of the same sentence, we can infer that, because it did not recognize some proper nouns, the algorithm was not able to discern whether *Thiago* was a simple or compound subject, neither diagnose an inflection problem verb t*o be*. Yet, in version 3, the proper noun was replaced by the personal pronoun *he*, evidence that the participant used the substitution strategy for supposing the non-recognition of a Brazilian proper noun. Thus, this replacement helped the algorithm clearly identify a simple subject, making the inflection *were* incorrect in the sentence.

One of the substitutions caused Jessica's text level to drop, visible in the chart entry in version number 3. This observation brings a reflection about the program's *modus operandi*, at least for beginner level constructions: the algorithm focuses on verbal agreement problems at the expense of vocabulary choice problems. We can affirm that because the construction *"He was drinking beer [...]"* obtained a feedback pointing to a serious error as opposed to *"Thiago were drinking beer"* or *"He was making a caipirinha [...]"* sentences that received the same type of feedback, pointing to a less serious sentence-level problem.

After considering all the strategies used by Jessica evidenced in the four versions of her text, we must consider that the learner's interaction with the application was beneficial, in the sense that it triggered the use of different written communication strategies, which provides a rich environment to enhance the skill being exercised. There were times when feedback caused uncertainty, providing unclear and unspecific feedback, especially when it was limited to orange-colored markings, causing the text proficiency to drop. Despite the adversities, from a global perspective, the learner improved her writing independently, assisted by the application.

It is worth mentioning that if she had been previously instructed on which strategies would help her write better, her rewritings could have become more elaborate. In any case, the autonomy presented by the beginner level participant interacting with a tutoring system fully formatted in the target language demonstrated an efficient interactive possibility. With the automated feedback support, the learner mobilized knowledge of the language she was learning to solve the linguistic problems pointed out by the system in an independent way. Such conduct has led to an increase in her writing proficiency level, *Write and Improve* major goal, even though developers have alerted that novice learners may not benefit as much.

During the writing workshop, Jessica dedicated herself to a particular task – *Life Goals and Dreams*. Possibly because of a topic that interested her, the participant wrote the text 7 times, the first 4 being rewritten without the interference of the teacher. The learner focused primarily on spelling problems and vocabulary choice directly pointed out by the feedback. After solving these problems, Jessica focused on correcting less obvious inadequacies, signaled by the *suspicious word* message, which also referred to a spelling problem, but the feedback could not provide tips. Then, the student tried to restructure her sentences with orange markings. However, from that

moment on, Jessica began to request the teacher's help because she was not able to diagnose other problems on her own.

Throughout the writing of this composition, a very common Portuguese speakers' error was perceived: the use of the verb *intend* to express future pretensions. However, this verb is a false cognate since it is only orthographically similar to the verb *pretender* from Portuguese, having very different meanings. For this inadequacy, the algorithm offered the solution directly, as can be seen in figure 57:

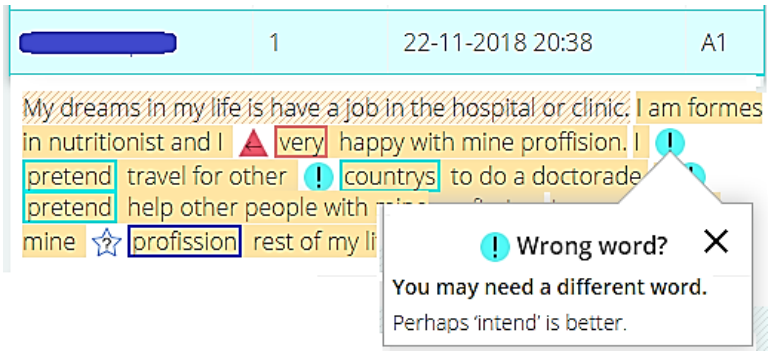


Figure 57 - Jéssica's text on November 22nd, 2018

When there is an error with a great incidence in *Write and Improve*'s database, it is able to recognize it without ambiguity and thus offer very specific feedback as the suggestion of replacement of the verb *pretend* by the verb *intend*. This is a very common mistake among English learners who speak Portuguese as ML, originated from the mental process of L1 transference, as described by Selinker (1972), as one of the processes that permeate the development of the subject's interlanguage. The orthographic similarity leads to the transfer of the ML language inventory term to the learner FL inventory. With the help of the system's feedback, Jessica realized this non-equivalence and accepted the suggestion offered, making use of the substitution strategy (Figure 58).

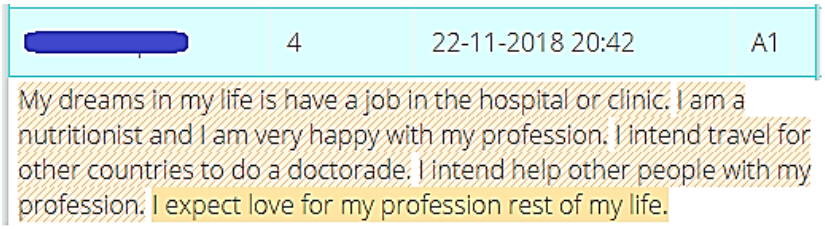


Figure 58 - Jéssica’s text after adjustment on November 22nd, 2018

The learner might have never perceived this problem in view of the process involved in the construction of her sentences being closely connected to her ML. Feedback with this character is extremely relevant considering that the application is intended to be used autonomously. From the graph shown in figure 59, we can see clearly how much Jessica’s interaction with the program was efficient once the level of her text *Life Goals and Dreams* presents an evolution from version 1 to version 4, where it obtained the best score. In the following versions, the teacher assisted the learner, guiding her in her sentences reconstructions and the graph indicates that the teacher's help was not as effective as the application in some moments, in view of the level drop in version 5. The improvement of Jessica’s text occurred more intensely during the first rewritings and in a more moderate way in the following versions.

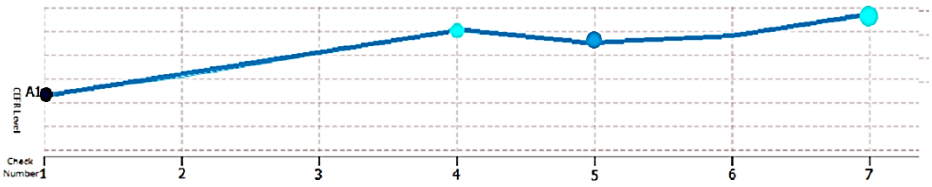


Figure 59 - Jéssica’s progress graph for one of her texts

In view of the above, the perception that human feedback always supplants automated feedback can be debated. Due to a set of inherent aspects concerning how the human mind processes information, we are led to believe in the improbability of an inanimate binary construct to overcome the complexity of neural connections to support such unpredictable dynamics as language learning. Still, during the writing workshop, the teacher sometimes felt a disadvantage, at a first glance simple, of the visual apparatus to point out problems in students' writing. Because the workshop occurred in the target language, that is, the use of Portuguese was not part of the work being developed, the lack of more objective signaling to show the students where the problems were and thus facilitate the activation of strategies, has hampered the effectiveness of some moments of oral feedback.

The teacher-student interaction was not so satisfactory since the communication had gaps that made it difficult to construct learner knowledge. Oral

feedback sometimes became a direct correction, in which the teacher pointed to the computer screen and said the correct word that the learner should write in a certain position. In the guidelines established by Shute (2007) for effective feedback, we recognize that this dynamic did little to build knowledge because the feedback was poorly elaborated. The immediate marking the system offers has become a better support to bring the users' attention to a particular element of their text. The tips that the program offers led to a moment of reflection on the learner's part, which did not seem to happen in the presence of the teacher, either because of a difficulty in understanding the oral feedback or by seeing in the teacher a shortcut to reach the final product. Such insight meets the discussion of Kluger and DeNisi (1996), once computerized and human feedback have distinct effects on learners, especially with respect to the sense of impartiality invoked by the machine over the subjectivity involuntarily manifested in the educator.

Another error learners frequently commit and therefore with a great occurrence in *Write and Improve's* database lies on the construction of the present simple and the present progressive tense, especially with regard to the use of the verb *to be*. In the investigation of the texts of the participant Jessica, we observed that the student omitted the verb *to be* in some of her constructions and every time this problem occurred, the algorithm detected it and provided adequate feedback. However, the hint is quite straightforward, suggesting exactly the verbal inflection needed to fix the sentence, as well as presenting two feedback message boxes for the same case. Still, this heightened attention to error can have a positive impact on the learner and trigger the acquisition of the structure. See Figure 60, where the feedback is illustrated:

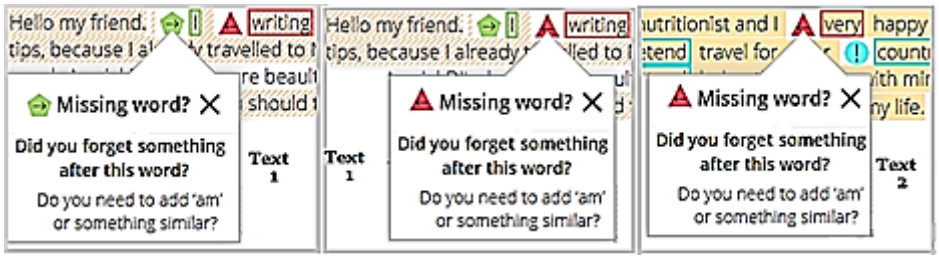


Figure 60 - Feedback for the omission of the verb *to be*

According to figure 60, in the excerpt entitled "text 1", we observe that the double indication of error occurs when the sentence constructed tries to form present

progressive time – *I am writing* – and the feedback points out that there is a word missing after the subject of sentence (*I*), indicated by a pentagon-shaped symbol in green color containing an arrow pointing to the right. There is also a marking before the main verb (*writing*), indicated by a red triangle-shape symbol containing an arrow pointing to the left.

In the feedback to the extreme right of the figure, in the excerpt titled "text 2", we note that there is only one feedback message box, with the green symbology, the sentence being constituted with the verb *to be – I am very happy*. As with the *pretend* verb case discussed earlier, this is a problem with a high incidence among Portuguese-speakers learning of English, which makes feedback very pertinent to address such an error. However, there is the observation on the organization of feedback for the two cases presented in figure 60, one with insufficient construction of the present progressive and the other with the omission of the main verb of the sentence. It is not possible to recognize the underlying reason for the second case also not to get a double marking and what the real pedagogical necessity of double marking the same error with different symbology. It is worth again the questioning of the language that was established in the program and its effectiveness in informing learners how they should solve the problem.

In order to understand how the feedback construction is done, starting from the error annotation and its subsequent coding to feed the algorithm, we presume that the same error occurs twice because of the way that it was transformed into code. The algorithm repeatedly points to inconsistency when the error has two distinct sources, and thus two codes, which makes the feedback unclear and misleading. It follows from this case that there is a refinement still necessary in error annotation – a fully human process –, the way RASP encodes the linguistic segments and builds the syntactic trees as well as the values assigned to the errors types, an aspect that influences directly in establishing the text proficiency level.

Another common incidence in Jessica’s texts was the irregularity in the verbal inflections, as the case of the use of the gerund to express past. The teacher realized that this inadequacy of her interlanguage took place after taking lessons on the perfect progressive present. Thus, the learner interlanguage undergoes a destabilization when this new content was inserted in her linguistic repertoire and her constructions began to oscillate between inflections of the verb in the simple past tense, in the perfect progressive tense and inflections containing gerund.

In one of her texts, illustrated in figure 55, we noticed that there is a verbal inflection that the learner did not use and did not obtain feedback. The algorithm only provided feedback for one of the verbs present in the composition – *talking*. The hint for this problem is to inflect the verb in the simple present tense (Figure 61).

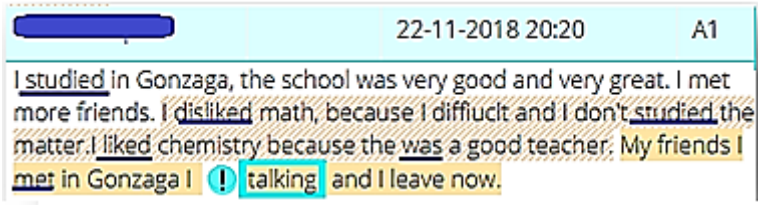


Figure 61 - Jessica's text with incorrect verb inflection on November 22nd, 2018

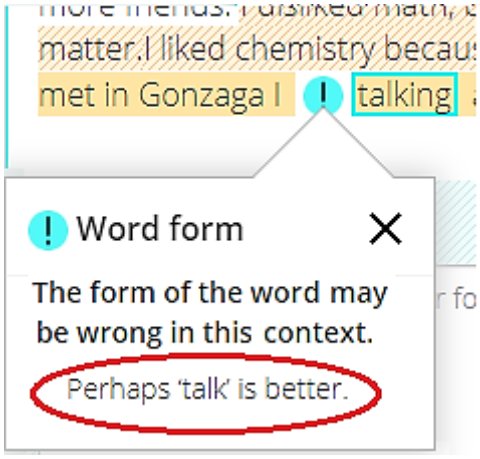


Figure 62 - Feedback to solve the incorrect verb inflection on November 22nd, 2018

Jessica complies with the suggestion and corrects the verb, however, it can not be determined whether the participant understood what in fact was inappropriate. The tip says that "*the form of the word seems to be wrong in this context. Maybe talk is better "* (Figure 62). The program suggests that the verb inflection be changed to the present tense, probably by the influence of the presence of the same inflection in another verb of the sentence – *leave* –, and a deictic element – *now*.

With respect to the tip in the feedback, the heading of the feedback box contains the message *Word form*, which suggests that the word form is incorrect. However, the form may be related to the word spelling, referring to its morphology. In the case quoted, the word, which is a verb, contains incorrect inflection, which does not refer to

its spelling, since *talking* is spelled correctly; it is the inflection that is inappropriate. This put, from a constructivist point of view, the provision of this feedback could cause more effect if these details were included in the message. For example, the message heading could bring the message *Verb form* instead of *Word form*, drawing the learners' attention to word class distinctions. The message below the heading could also bring the question of the form of the verb – not the word – being wrong. It may seem a very superficial debate; however, small details as part of a whole learning process can help or tamper with learners constructions of hypothesis about the language, especially if they are to use the application without human assistance.

In addition, the message comments on the word being wrong in that context, which does not comply with the truth because the program is not context-sensitive at that level, it lacks the ability to assess the composition context to judge whether a term is in accordance with its surroundings, a feature already discussed at the beginning of this chapter. The sensitivity of the program is related to the question of whether what learners wrote is in line with the theme of the task they chose, which had a good score in Jessica's case (Figure 63).

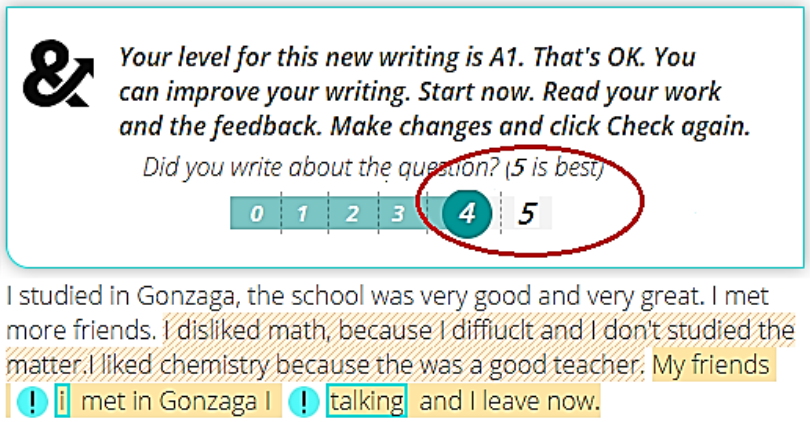


Figure 63 - Overview feedback on Jessica's text on November 22nd, 2018

Considering Jessica's difficulty at using the past tense of verbs and inflecting accurately the verb *to be*, peculiarities verified by the teacher's observation to her texts, the application was not effective in contributing to the improvement of this aspect of her interlanguage. The application's capacity relies on evaluating sentences and word sets in isolation, regardless of the volume of similar errors in the text as a whole os in all the texts the learners submit. *Write and Improve* is not able to record the incidence

of errors from the same origin of a single user to customize ways in which learners can overcome their difficulties and improve their linguistic competence. Thus, the issue of adaptability of the application to individual needs in autonomous learning is not yet a reality, given the limitations of this nature. On the other hand, if artificial intelligence does not reach this level of assistance to learners, it certainly offers tools that facilitate observation of learners' development by the teacher/tutor/instructor responsible for guiding the process, facilitating the feasibility of teaching individualization.

7 Challenges and Suggestions

Automated tutoring systems for EFL writing detain a great potential in providing useful feedback. As we have pointed out in chapter 2, there has been few studies analyzing the implementation of such technology to assist learners' writing development. Besides, there are many aspects to be noted in order to make these intelligent tutors useful assistants. That being said, it is important to observe how much their feedback triggers learners' mental processes, because that is paramount to guarantee high chances for them to offer efficient help, leading to written language development.

One of the main aspects analyzed in *Write and Improve* was its feedback taxonomy. It consists of simplified messages in order to reach learners since beginner level. However, feedback simplification may not be sufficient to cause learners to make use of more effective strategies, making their rewriting attempts tiresome sometimes. There are only 5 symbols used to discriminate the problems in the texts and, considering the wide range of possibilities language can generate, 5 is a small number. Although each of these 5 codes can yield different feedback messages, only 10 different messages could be observed to the subjects' texts during the experiment.

The program developers state that there are 80 types of errors coded into the system. However, there is not a way to know what these codes linguistically account for unless the mistakes they are supposed to correct happen in a text submitted to feedback. Taking this into account, if teachers were able to retrieve this information beforehand, especially inside the *+ClassView* utility, it would be certainly helpful. They would know in advance what problems the system targets in the texts and, therefore, they could instruct learners more specifically about what the feedback messages mean as well as how to approach them.

When feedback triggers written communication strategies effectively, there is room for learning to happen. Among the strategies observed in the subjects' rewritings, restructuring was a recurrent one. It is a valid strategy and certainly plays a role in writing skill development. However, its overuse may not be the most beneficial way for learning to be longstanding. Taking Daniel's endeavor of 41 rewritings of a text and his proficiency level fluctuation, we could see such an effort was not worth it once his proficiency level did not go further than C1. Besides that, within 20 of those restructurings and submissions, his level kept at B2.

Having said that, it would be interesting if the system could somehow detect areas or sentences that learners keep insistently restructuring and provide a different feedback. As a suggestion, maybe the system could signal that the sentences being frequently changed would not cause the text's proficiency level to rise or drop significantly. A more specific feedback on the problematic sentences themselves, on a word level, would be very helpful. This way, students would see an approach to their writing different from the unspecific colored marking, which is one of the triggers to the use of restructuring. This kind of procedure would mean perfecting the tool from the users' and teachers' perspective and needs and it would also improve the interactivity patterns.

As we have seen in the data analysis, the substitution strategy can be triggered satisfactorily once the error pointed out by the feedback relates more closely to learners' ML. Nonetheless, it is important to mention that in the case analyzed in this research, the subject managed to make the right correction certainly because the feedback also brought an option for the substitution. In order to account for this, a promising as well as challenging approach would be relate feedback tips to mental processes such as ML transference, which constantly permeates learners' assumptions about how language works, especially in beginner levels. This way, the feedback could be somehow linked to a set of common transference errors regarding users' ML. That would probably demand a substantial reorganization in the data collection and in the arrangement of error codes to distinguish the errors among languages. In addition, the message that is provided to learners could be less direct, as not showing the exact term to substitute another, but showing the way for leaners to get there by themselves. It suffices to say that not all transference errors will be triggered successfully, because there are many other variables influencing the process. It could be an astute leap in raising the system's level of effectiveness.

When the user does not understand well what a message tries to convey, for example, some mental processes are not triggered. Consequently, communication strategies may not be activated and therefore learning may not happen. It is relevant to point out that these events are not necessarily sequential and isolated from external stimuli that interact in concomitance with reasoning. From this perspective, it seems reasonable to attempt a design based on FL acquisition assumptions. A good communication channel is a valuable asset to provide good assistance to learners. Without this foundation, one can easily compromise the triggering of mental processes

and the tool can eventually become irrelevant regardless of interface, ease of access, gratuity, etc. In order to pursue a functional hybrid instructional environment, aspects such as communication, interaction and expectations – both from teachers and learners – need to be well adjusted. Otherwise, much time and effort might be consumed and a feasible outcome may never happen.

Write and Improve uses the colors green, red and blue to point out inadequacies in learners' texts. The color green has as broad meaning the idea of permitted, correct and approved. When this specific color is used to refer to an error, something prohibited, incorrect, and disapproved, there is a high possibility that the message is misinterpreted and communication fails. In addition, it is relevant to note that, in the case presented in the analysis, the same error received different color feedback. Furthermore, the positioning of the symbology did not indicate where there was a missing element, which made it inexact and little useful. Even without undertaking in depth in the field of semiotics, we could observe that the learner-program interaction did not happen in a satisfactory way under some conditions, causing learners to hesitate on which path to take. From a teacher's perspective, the most suitable strategy to be triggered in order to solve the inadequacy appropriately was restructuring, in which a new verbal plan is made for the sentence. However, the learner used the substitution strategy, replacing an element that was also necessary due to feedback misunderstanding. Once this observation has been made, double marking the way it was presented did not favor effective communication and may compromise the construction of language knowledge by learners.

Problems of collocations formation, spelling and use of determinants are easily identified by the algorithm, and always make part of the initial feedback. When these problems are properly adjusted, the text's proficiency level usually rises. On the other hand, some syntax problems such as anaphora and gerund as sentence subject and object are covered only by sentence-level feedback (orange marking), or are not detected at all, which guides the learner very vaguely. These findings also indicate a difficulty of natural language processing by the algorithm used in *Write and Improve*. Since anaphora is a device employed to resume elements mentioned before in a text and the system is not able to analyze pieces of text larger than a sentence, all terms used to make reference to something mentioned in a previous moment are not identified. This can eventually lead to inadequate feedback and consequently unnecessary correction from learners. Working on a way to make the algorithm capture

this structural elements would bring a relevant contribution to the feedback precision and effectiveness.

The aspect of rewriting is one of the cornerstones of the application, once texts can be altered and submitted to feedback infinite times. This way of working with texts inspires the fruitful idea that developing the writing skill is a continuous process of rebuilding language. However, we could observe, through the graphs the system provides, a certain pattern in the level oscillation, which happened every 5 submissions or so. That means texts had a rising level up to the fifth version, and this level tended to drop again in the following versions. This could be noticed in texts that had from 5 versions on. This might be an indication of how the algorithm works, following a pre-established assessment pattern. In Eco's (1991) perspective, we should not be surprised since computer programs work with combinatory possibilities, not with meaning and semantics. Language semantics plays an irrefutable role in producing meaning and enabling communication. Here lies, indeed, a substantial challenge for tutoring systems to improve.

From this finding, we can think about how unbeneficial it can be to have so many submissions to feedback. Learners might feel frustrated rewriting the texts so many times and not perceiving a real progress in their proficiency level. It is important to mention that, in this patterned fluctuation, the text level never exceeded the highest level reached in the subject's first five submissions, making excessive rewritings somehow unproductive from a formative assessment perspective. Certainly, the rewritings have brought better sentence constructions and spelling that is more accurate. Nevertheless, learners as well as teachers who figure this apparent evaluation pattern out in their writings might mistrust the precision of the system's assessment capability and consequently change their approach to using the application.

The way feedback appeared on the screen in terms of language and in visual aids was imperative for communication to be effective, because the users were, in the first place, language learners. That aspect also affected the system's credibility, once learners expected the teacher's final comment on their texts to revise the automated feedback work. Learners felt a discrepancy between how the system approached the text inadequacies and how the teacher used to work their errors with them. That being said, it seems pertinent to bring the language and approach of the system closer to that of a teacher, beyond dialogue-like motivational messages and tips. There was a

demand for teacher validation before learners could follow the system's correction confidently. The machine needs to learn the language of the human and the human, in turn, the language of the machine. This can fulfill pedagogical as well as psychological needs that might be crucial for tutoring systems use in large scale to prosper and be normalized in educational praxis. In the words of Lippman[20] (1988 apud Primo, 1998), "instead of working with the idea of a human-machine relation, consider it a human-human relation."

Knowledge exchange takes place through language comprehension. For that, the more the users understand the language used by the program, the more likely they are able to improve their skills. Even though the use of English to establish communication with learners is visible on the surface, it is the underlying language that needs to be in line with the capacity for human understanding. The binary texture that translates into language needs good planning to be understood by learners.

We have analyzed the learner-program relationship and the triad that includes the teacher in this dynamics fomented by *Write and Improve*. We could sense a tension, however subtle, in being in the presence of a teacher/tutor/instructor doing the assessment of a written production. This discomfort does not exist or at least is minimized when feedback is presented without the human element or even when human feedback is sent remotely. Questions of judgment, so present when dealing with adults, as they are exposing their capacities and limitations in the teaching-learning environment, seem to be mitigated by the neutrality of the machine. The impersonation of the feedback process softens the fear of a reputation hurt by the master and assists in the work of each learner's affective filter. That being said, making use of automated feedback to correct textual irregularities can prove to be of excellent assistance in terms of lowering learners' anxiety and building more self-confidence in the learning process.

In the studies brought to build the foundation of this work, a common approach to the use of automated writing assessors was adopted. Researchers previously instructed the subjects and teachers about how the tool worked before starting the experiment for data collection. The approach to *Write and Improve* was different. We had not showed any features of the system beforehand because we aspired to see how the human-machine interactions would arise without any interference. This

[20] Our translation. In the original: "Em vez de trabalhar com a ideia de relacionamento entre homens e máquinas, considere pessoas com pessoas"

approach was thought taking into consideration the autonomy aspect that is very closely linked to the use of assistive learning technologies. Since this application is meant to be used by learners with or without the assistance of a teacher/tutor/instructor, we desired to see how intuitive its features were specially being all in the TL. Our assumptions were right. The communication between learners and the program would have been more effective if they had been instructed before and their expectations had been adjusted. Therefore, at least regarding this specific writing assessor, it is not advisable to entrust all the potential of the automated tutor to its level of interactivity. There must be human involvement even if in an introductory moment to ensure the communication will be the most efficient it can be and interactions can afford rich learning moments in an autonomous process.

Write and Improve developers refer to their application's language system as Learner English, which means their database is fed by texts produced by English learners, nonnative speakers, which have their own peculiar and individual developing interlanguage. Nevertheless, the feedback provided by the system for spelling problems is solely based on standard British English. Thus, the feedback points out any forms that do not represent this variant. This limitation ignores any other language source from which learners may have had contact with and which helped build up their interlanguage. Taking this into account, it is prudent to consider having specialists who acknowledge these differences when coding errors and feeding the system to provide feedback. Adding to the team linguists who are speakers of other variants themselves seems also desirable, measure that can help build a system with an embracing *modus operandi,* capable of detecting not only British English. English as a world language, a lingua franca, cannot be disregarded if educational technology companies seek to have popular and efficient products. Correcting such differences not only brings confusion and frustration to the learners as well as does not contribute to the application's credibility.

Therefore, we verify that it is still a challenge in computational linguistics to account for the volume of variables that arise when dealing with such a dynamic phenomenon as language is. Natural language processing executed by complex algorithms still does not equate to the refinement of the human mind to correct linguistic errors. A teacher is able to evaluate a variety of concomitant aspects and correlate them in a matter of seconds to correct errors in a student's writing. A computer program has restrictions that teachers easily use in their favor, such as characterize the subject

and relate their productions in a global and personalized way simultaneously. Human reading goes beyond the surface of the text and expands over a vast area that includes prior knowledge, experiences, personality traits, learning style, weaknesses and strengths in language development, to mention a few.

Undoubtedly, we are posthuman beings, in the sense we can attain further capacities by adopting new technological means. Taking the automated feedback by *Write and Improve* as an example, we are able to have 200-word length texts corrected in around 15 seconds. No human brain can do that so far. From one perspective, these advancements sound amusing, however, as Leonhard (2016) comments on technology's incredible speed but still crude discernment:

> "a super-computer can win in chess or GO but can currently not talk to a 2-year old. A person that meets me in a hallway somewhere needs an average of 1.4 seconds to gain some kind of basic understanding about me, even without speaking – a computer still does not really understand my values and feelings after it has ingested my entire browsing and social network history of the past 7 years (an estimated 200 Million data points)."

Thus, we must have clarity and apply technology in enterprises that algorithms excel at, which include working with probability and programming, statistics and mapping. Unfortunately or not, language use holds much more than this and, for that matter, we should work with AI to the extent it can do a good job. Recognizing its limitations to deal with creativity, unpredictability and retrieval of elements that are not possible to be coded, such as intentionality, empathy and irony, means we should certainly leave some of the work to human care. On the other hand, in order to do that, we human teachers must let go of the thirst for power and control that sometimes manifests on us and distribute the workload among non-human peers, benefitting from their features that best suit our pedagogical needs.

The symbiosis between human and algorithmic cognition through intelligent language learning systems can, in the long run, result in a mixed interlanguage, through mutual adaptations between human and machine, incorporating traits from both organic and artificial, with interactions in favor of new possible communicative demands in a posthuman scenario of relationships. The standard language referenced by native speakers since the beginning of times and regarded as an emblem of cultural and geographical power and mastery, from a posthuman perspective, will be traversed by the artificial language processed by intelligent systems that will invariably cause structural deviations that will eventually belong to common use. Just as the physical

and time boundaries have been diluted with the advent of the World Wide Web, the posthuman language will also be the result of this erasure allied with the cybernetic agency that permeates social relations.

We will now have one more agent participating in the inherent transformation in language. Humans will share with artificial intelligence the task of manipulating language to meet communicative needs. The human mind leaves the prominent position in language processing to adapt to a context populated by humans and nonhumans capable of communicating through verbal language. Understanding the situation in the years to come is paramount to establishing a fruitful environment for language teaching and learning. Technology will not cease to advance and occupy previously human spaces and that it is not a gold rush to see who succeeds with perfection. Quite the opposite, by the time we face this reality and think about the benefits that this intertwinement can bring we are going to understand the need to develop new skills as well as innovate our pedagogical references to deal with education assisted by intelligent tutoring technologies.

Posthuman applied linguistics in turn must endeavor to bring these issues to the fore, as much is advocated for a more computerized pedagogy to keep up with technological developments, disperse the teacher's workload and make the process generally more practical. However, we are uncertain of the ways in which embracing this condition may lead us. Breaking with the view of the human as the center of our existence causes insecurity, as presented in the analysis of this research, in which the learner only gave credit to the system when the teacher gave her own endorsement of the feedback. As Chislenko (1995) brings in his work, we have become fyborgs – functional cyborgs, in which our bodies get cybernetic extensions, such as our mobile devices (E-readers, cell phones, tablets, laptops), and from there we are able to perform new and sometimes more complex tasks than before. We already trust computer systems to memorize our many passwords and organize our phone contacts and computer files to such degree that there will not be long until tutoring systems for language learning are normalized in our teaching scenarios.

The next pages should be written by rethinking our position as thinking beings and our capacity to construct a paradigm consonant with the decentralization of the subject and opening the stage for the agency of the inanimate also as a producer of knowledge and capable of handling various tasks. It is well known that it is not human nature to share power and yield spaces. However, it is indispensable to view this

transition never as a dispute, but as a synchronization of entities. The same way Haraway (1986) brings the conception of cyborg in her manifesto, we are heading towards mergers, where organic and machinic intertwine and give rise to new configurations of being and doing. Humankind has created artifacts from the earliest times with the intention of facilitating the endeavor of figuring the world out and consequently needs to transform itself to deal with its own creations. In posthuman thought, this dynamics shift from a single way flow, in which creature serves or adapts to the creator, but a two-way relation, where humans also adapt and shape themselves to their own creation without fearing the loss of control or identity. In the words of Hayles (1999, p. 287) “[…] rather than disembodied information, the posthuman offers resources for rethinking the articulation of humans with intelligent machines”, meaning that in order to thrive in the decades to come we must embrace technology and eventually transfigure teaching and learning current paradigms.

References

ALESSI, S. M.; TROLLIP, S. R. **Multimedia for learning**: Methods and development. Boston: Allyn and Bacon, 2001.

ALIAKBARI, M.; KARIMI ALLVAR, N. Communication strategies in the written medium: The effect of language proficiency. **Linguistik Online**, v. 40, n. 4, p. 3-13, 2009. Retrieved from https://bop.unibe.ch/linguistik-online/article/view/428/687 on Oct 08, 2019.

ÅSTRÖM, K. J.; MURRAY, R. M. **Feedback Systems**: An Introduction for Scientists and Engineers. New Jersey: Princeton University Press, 2008.

ATTALI, Y.; BURSTEIN, J. **Automated Essay Scoring with E-rater** v.2.0. RR-04-45, Educational Testing Service. Princeton: New Jersey, 2006. Retrieved from https://www.ets.org/Media/Research/pdf/RR-04-45.pdf on Apr 10, 2018.

BARABASI, A. L. **Linked**: the new science of networks. New York: Plume, 2002.

_____. **Bursts**: the hidden pattern behind everything we do, from our emails to bloody crusades. New York: Dutton, 2010.

BARONE, C.A. The changing landscape and the new academy. **Educause Review**, v. 38, n. 5, p.41-47, 2003. Retrieved from http://net.educause.edu/ir/library/pdf/ERM0353.pdf on 14 Jul 2018.

BATES, T. **Teaching in a Digital Age**: Guidelines for designing teaching and learning. [ebook], Tony Bates Associates Ltd, 2015. Retrieved from https://opentextbc.ca/teachinginadigitalage/ on 28 Feb 2019.

BELL, B. S.; FEDERMAN, J. E. E-learning in postsecondary education. **The Future of Children**, Cornell University: School of Industrial and Labor, v. 23, n. 1, p. 165-185, 2013. Retrieved from http://digitalcommons.ilr.cornell.edu/articles/928 on Jun 19, 2018.

BERTALANFFY, L. V. **Teoria geral dos sistemas**. Petrópolis: Vozes, 1973.

BETSY. (s.d.). Retrieved from http://edres.org/betsy on Apr 28, 2018.

BLOMMAERT, J. Meaning as a nonlinear effect: The birth of cool. **Tilburg Papers in Culture Studies**, p. 1-21, 2014.

BLOOMFIELD, L. **Language**. Londres: George Allen and Unwin, 1933.

BRAUN, K. F. **Electrical Oscillations and Wireless Telegraphy**. Nobel Lecture, 11 Dec. 1909. Retrieved from https://en.wikipedia.org/wiki/Feedback#cite_note-7 on 07 Sep 2017.

BRIGGS, J. **Fractals**: The patterns of chaos. New York: Simon and Schuster, 1992.

BRISCOE, T. The use of lexical resources in Cambridge Write & Improve. In: KOSEM, I., JAKUBÍČEK, M., KALLAS, J., KREK, S. (eds.). Electronic lexicography in the 21st century: linking lexical data in the digital age. **Proceedings of the eLex 2015 conference**. Herstmonceux Castle, United Kingdom. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd., 2015. Retrieved from https://elex.link/elex2015/videos/ on Jul 18, 2018.

BROPHY, J. E. Teacher praise: A functional analysis. **Review of Educational Research**, v. 51, n. 1, p. 5-32, 1981.

BROWN, J.S. Growing up digital: how the web changes work, education, and the way people learn. **USDLA Journal**, v. 16, n. 2, 2002. Retrieved from www.johnseelybrown.com/Growing_up_digital.pdf on Jul 14, 2018.

CANAGARAJAH, S. Lingua Franca English, Multilingual Communities, and Language Acquisition, **The Modern Language Journal**, n. 91, p. 923-939, 2007.

CARDOSO, A. C. S. Feedback em contextos de ensino-aprendizagem on-line. **Linguagens e Diálogos**, v. 2, n. 2, p. 17-34, 2011.

CHEN, E.; CHENG, E. Beyond the Design of Automated Writing Evaluation: Pedagogical Practices and Perceived Learning Effectiveness in EFL Writing Classes. **Language Learning and Technology**, p. 94-112, 2008. Retrieved from http://llt.msu.edu/vol12num2/chencheng.pdf on Feb 4, 2018.

CHIMBGANDA, A. B. Communication strategies used in the writing of answers in biology by ESL first year science students of the University of Botswana. **English for Specific Purposes**, v.19, n.14, p. 305-329, 2000.

CHISLENKO, A. **Legacy systems and Functional cyborgization of humans**. [online].1995. Retrieved from: http://www.lucifer.com/~sasha/articles/cyborgs.html on Jan 22, 2020.

CHOU H. C.; MOSLEHPOUR, M.; YANG, C.Y. My Access and Writing Error Corrections of EFL College Pre-intermediate Students. **International Journal of Education**, v. 8, n. 1, p. 144-161, 2016.

COUNCIL OF EUROPE. **Common European Framework of Reference for Languages**: learning, teaching, assessment. Cambridge: Cambridge University Press, 2001. Retrieved from https://rm.coe.int/16802fc1bf on Oct 05, 2019.

DATA POPULAR INSTITUTE. *Learning English in Brazil*: Understanding the aims and expectations of the Brazilian emerging middle classes. 1st ed. São Paulo: British Council, p. 5-29, 2014. Retrieved from https://englishagenda.britshcouncil.org/sites/default/files/attachments/estudov13_englishinbrazil.pdf on Nov 17, 2018.

DAVIES, G. Introduction to multimedia CALL. Module 2.2. *In:* DAVIES, G. (ed.) **Information and Communications Technology for Language Teachers**

(ICT4LT), Slough, Thames Valley University [Online], 2012. Retrieved from http://www.ict4lt.org/en/en_mod2-2.htm on Dec 29, 2018.

DEMETRIADIS, S.; CADOZ, C. A conceptual framework for the integrative design of adaptable representations for learning, In: Mendez-Vilas, A.; Gonzalez-Pereira, B.; J. Mesa Gonzalez, J.; Mesa Gonzalez, J. (Eds.). **Recent Research Developments in Learning Technologies**, Badajoz, Spain: Formatex, 2005, p. 1250-1255.

DICKINSON, D. Multimedia myths. **Australian Personal Computer**, v. 16, n. 10, p. 144-145, 1995.

DIKLI, S. An Overview of Automated Scoring of Essays. **The Journal of Technology, Learning, and Assessment**, v. 5, n. 1, p. 5-35, 2006. Retrieved from www.jtla.org on Apr 7, 2018.

DÖRNYEI, Z. On the teachability of communication strategies. **TESOL Quarterly**, v. 29, p. 55-85, 1995.

_____.; SCOTT, M. L. Communication Strategies in a Second Language: Definitions and Taxonomies. **Language Learning**. v. 47, n.1, p. 173-210, 1997.

_____, J.; CHENG, V. A Toolkit to Assist L2 Learners Become Independent Writers. **Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics and Writing**. Los Angeles, California, p. 33-41, 2010.

DOWNES, E. J.; MCMILLAN, S. J. Defining interactivity: a qualitative identification of key dimensions. **New Media Society**, v. 2, n. 2, p. 157-179, 2000. Retrieved from https://www.researchgate.net/publication/245852959_Defining_Interactivity_A_Qualitative_Identification_of_Key_Dimensions on Aug 18, 2018.

DOWNES, S. **An introduction to connective knowledge**. [online]. 22 dezembro 2005. Retrieved from https://www.downes.ca/cgi-bin/page.cgi?post=33034 on Jul 28, 2018.

DUARTE, G. B. **Eventos complexos de letramentos na aprendizagem de inglês:** relações entre práticas de letramentos, gamificação e motivação. 2017. Tese (Doutorado em Linguística Aplicada) – Universidade Católica de Pelotas, Pelotas, 2017.

DULAY, H.; BURT, M.; KRASHEN, S. **Language Two**. New York: Oxford University Press, 1982.

ECO, U. **Tratado geral de semiótica**. 2a ed. São Paulo: Perspectiva, 1991.

EDUCATIONAL TESTING SERVICE (ETS). (s.d.). Retrieved from http://www.ets.org on Apr 12, 2018.

ELLIS, R.; BARKHUZEIN, G. **Analyzing Learner Language**. Oxford University Press, 2005.

_____. A typology of written corrective feedback types. **ELT Journal**, v. 63, n. 2, p. 97-107, 2009.

FÆRCH, C., & KASPER, G. Plans and strategies in foreign language communication. In: FÆRCH, C.; KASPER, G (Eds.). **Strategies in interlanguage communication**. Harlow, UK: Longman, p. 20-60, 1983b.

FERREIRA, H. C. A teoria piagetiana da equilibração e as suas consequências educacionais. **Série Estudos**. Instituto Politécnico de Bragança, v. 55, p. 7-26, 2003. Retrieved from https://bibliotecadigital.ipb.pt/bitstream/10198/208/1/55%20-%20A%20teoria%20piagetiana%20da%20equilibra%C3%A7%C3%A3o%20e%20as%20suas%20consequ%C3%AAncias%20educacionais.pdf on Jul 19, 2018.

FIALHO, V. R. **Comunidades virtuais na formação de professores de espanhol língua estrangeira a distância na perspectiva da complexidade**. 2011. Tese (Doutorado em Linguística Aplicada) – Universidade Católica de Pelotas, Pelotas, 2011.

FILATRO, A**. Design instrucional na prática**. São Paulo: Pearson Education do Brasil, 2008.

HARAWAY, D. A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century," In: HARAWAY, D. **Simians, Cyborgs and Women**: The Reinvention of Nature. New York: Routledge, 1991, p. 149-181.

HATTIE, J.; TIMPERLEY, H. The Power of Feedback, **Review of Educational Research**, University of Auckland, v. 77, n. 1, p. 81-112, 2007. Retrieved from http://education.qld.gov.au/staff/development/performance/resources/readings/power-feedback.pdf on Sept 1, 2018.

HIRSH-PASEK, K.; SCHLESINGER, M.; GOLINKOFF, M.; CARE, E. **The new-humanism**: the technology should enhance, not replace human interactions. [online]. 11 Junho 2018. Retrieved from https://www.brookings.edu/blog/education-plus-development/2018/06/11/the-new-humanism-technology-should-enhance-not-replace-human-interactions/ on Jul 31, 2018.

HUBBARD, P.; IOANNOU-GEORGIOU, S. Teaching English Reflectively with Technology. **IATEFL Learning Technologies Special Interest Group**, p. 4-33, 2017.

JENSENN, J. F. Interactivity: tracing a new concept in media and communication studies, **Nardicom Reviews**, v. 19, n. 1, p. 185-204, 1998.

KERR, P. **A challenge to connectivism**. [online]. December 26, 2006. Retrieved from http://billkerr2.blogspot.com/2006/12/challenge-to-connectivism.html on Jul 27, 2018.

KIELTY, L. Feedback in Distance Learning: Do students' perceptions of corrective feedback affect retention in Distance Learning? **Theses and Dissertations. Paper 1114**. University of South Florida, 2004.

KLUGER, A. N.; DENISI, A. The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. **Psychological Bulletin**, v. 119, n. 2, p. 254-284, 1996. Disponível: http://psycnet.apa.org/fulltext/1996-02773-003.pdf on Aug 28, 2018.

KOOLSTRA, C. M.; BOS, M. J. W. The development of an instrument to determine different levels of interactivity, **The International Communication Gazette**, v. 71, n.5, p. 373-391, 2009. Retrieved from https://research.vu.nl/ws/portalfiles/portal/2427796/219094.pdf on Aug 18, 2018.

KRAMSCH, C.; WHITESIDE, A. **What is symbolic competence and what can we do with it?** [online]. Lecture at Berkeley Language Center, University of California: Berkeley, California, 2015 Retrieved from http://blc.berkeley.edu/2015/08/19/lectures-by-c-kramsch-a-whiteside-november-20-2015/ on Oct 20, 2018.

KRASHEN, S. Comprehensible input and some competing hypotheses. In: COURCHÊNE, R. et al. **Comprehension-based Second Language Teaching**, Ottawa, Canada: University of Ottawa Press, 1992, p. 19-38.

LANTOLF, J. **Sociocultural Theory and Second Language Learning**. Oxford: Oxford University Press, 2000.

LARSEN-FREEMAN, D. Chaos/complexity science and second language acquisition. **Applied Linguistics**. Oxford, v. 2, n. 18, p.141-165, 1997.

LEFFA, V. J. Uma ferramenta de autoria para o professor: o que é e o que faz. **Letras de Hoje**, Porto Alegre, v. 41, n. 144, p. 189-214, 2006.

LEKI, I. Coaching from the margins: Issues in written response**.** In: KROLL, B. **Second Language Writing:** Research Insights for the Classroom. Cambridge: Cambridge University Press, 1990, p. 57- 68.

LEWIN, R. **Complexidade**: a vida no limite do caos. Rio de Janeiro: Rocco, 1994.

LÜDKE, M.; ANDRÉ, M.E.D.A. Pesquisa em educação: abordagens qualitativas. São Paulo: EPU, 1986.

MARTINS, E. E. B.; VIANA, R. F. Por uma visão de linguagem ciborgue e coletiva. **Trabalhos em lingúística Aplicada**, n. 58, v. 2, p. 496-519, 2019. Retrieved from https://periodicos.sbu.unicamp.br/ojs/index.php/tla/article/view/8655578 on Oct 15, 2019.

MASLOW, A. H. **The Farther Reaches of Human Nature**. New York: Penguin Books, 1971.

MASON, B.; BRUNING, R. Providing Feedback in Computer-based Instruction: What the Research tells us, **CLASS Project Research Report n. 9**, Center for Instructional Innovation, University of Nebraska-Lincoln, 2001.

MCNAMARA, D. S.; CROSSLEY, A. S.; ROSCOE, R. Natural language processing in an intelligent writing strategy tutoring system. **Behavior Research Methods**, n. 45, p. 499-515, 2013. Retrieved from https://link.springer.com/content/pdf/10.3758%2Fs13428-012-0258-1.pdf on Apr 24, 2018.

MILTON, J. Mark My Words: Technologies for supporting, managing and responding to student writing. **Proceedings of the Second Teaching and Learning Symposium**, Hong Kong, Senate Committee on Teaching and Learning Quality, and Center for Enhanced Learning and Teaching, HKUST, p. 1-8, 2004.

MOATTARIAN, A.; TAHRIRIAN, M. H. Communication Strategies Used in Oral and Written Performances of EFL Learners from Different Proficiency Levels: The Case of Iranian EFL University Students, **Journal of Foreign Language Teaching and Translation Studies**, v. 2, n. 1, p. 21-37, 2013. Retrieved from http://efl.shbu.ac.ir/article_79179_e915cea1a4e81a66a97a74f96a44a4d4.pdf on Oct 10, 2019.

MOHSEN, M. A.; ALSHAHRANI, A. The effectiveness of using a hybrid mode of automated writing evaluation system on EFL students' writing. **Teaching English with Technology**, n. 19, v. 1, p. 118-131, 2019. Retrieved from https://www.tewtjournal.org/issues/volume-2019/volume-19-issue-1/ on Nov 1, 2019.

MOORE, M. G. Three types of interaction. **American Journal of Distance Education**, v.3, n. 2, p. 1-7, 1989.

NAIDU, S. E-Learning: A Guidebook of Principles, Procedures and Practices. **Commonwealth of Learning**. New Delhi: Commonwealth Educational Media Center for Asia. 2ª Ed., 2006. Retrieved from http://oasis.col.org/bitstream/handle/11599/53/e-learning_guidebook.pdf?sequence=1&isAllowed=y on May 27, 2019.

NICOL, D. From monologue to dialogue: improving written feedback processes in mass higher education. **Assessment and Evaluation in Higher Education**, v. 35, n. 5, p. 501-517, 2010.

NUNES, G. M. **A escrita em inglês como sistema adaptativo complexo**: uma abordagem colaborativa na aprendizagem de língua estrangeira por meio das TIC. 2013. Dissertação (Mestrado em Linguística Aplicada e Estudos da Linguagem) – Universidade Federal de Pelotas, Pelotas, 2013.

O'LEARY, Z. **The Essential Guide to doing research**. London: SAGE Publications, 2004.

PAGE, E. B. Project Essay Grade: PEG. In: SHERMIS, M.; BURSTEIN, J. (Eds.). **Automated essay scoring***: A cross-disciplinary perspective.* Mahwah, NJ: Lawrence Erlbaum Associates, 2003, p. 43-54.

PAIVA, V.L.M.O. **Aquisição de segunda língua**. São Paulo: Parábola Editorial, 2014.

_____ . Como o sujeito vê a aquisição de segunda língua. In: CORTINA, A.; NASSER. S.M.G.C. **Sujeito e Linguagem**. São Paulo: Cultura Acadêmica, 2009. Retrieved from http://www.veramenezes.com/sujeito.pdf on Jan 25, 2019.

_____ . Feedback em ambiente virtual In: LEFFA.V. (Org.) **Interação na aprendizagem das línguas**. Pelotas: EDUCAT, 2003. p. 219-254.

_____. Linguagem e aquisição de segunda língua na perspectiva dos sistemas complexos. In: BURGO, V. H.; FERREIRA, E. F.; STORTO, L. J. **Análise de textos falados e escritos**: aplicando teorias. Curitiba: Editora CRV, 2011. p. 71-86.

_____ . Modelo fractal de aquisição de línguas. In: BRUNO, F. C. (Org.) **Reflexão e Prática em Ensino/Aprendizagem de Língua Estrangeira**. São Paulo: Editora Clara Luz, 2005, p. 23-36.

PENNYCOOK. A. **Posthumanist Applied Linguistics**. Routledge: Taylor & Francis Group, 2018.

PRENSKY, M. Digital natives, digital immigrants, Part I: A new way to look at ourselves and our kids. **On the Horizon**, MCB University Press, v. 9, n. 5, p. 1-6, 2001. Retrieved from http://www.marcprensky.com/writing/Prensky%20-%20Digital%20Natives,%20Digital%20Immigrants%20-%20Part1.pdf on Jul 14, 2018.

PYKE, J.; SHERLOCK, J. A closer look at instructor-student feedback online: a case study analysis of the types and frequency. **MERLOT Journal of Online Learning and Teaching**, v. 6, n. 1, 2010.

QUADROS, R. M; FINGER, I. **Teorias de Aquisição de Linguagem**. Florianópolis: Editora da USFC, 2008.

RAMAPRASAD, A. On the Definition of Feedback. **Systems Research and Behavioral Science**. John Wiley & Sons, Ltd, v. 28, n. 1, p. 4-13, 1983.

RINVOLUCRI, M. Key concepts in ELT: feedback. **ELT Journal**. v. 48, n. 3, p. 287-288, 1994.

ROGERS, C. R. The foundations of the person-centered approach. In: ROGERS, C. R. **A way of being**. Boston: Houghton Mifflin, 1980, p. 113-136.

ROSA, S. H. D. **O feedback oral do professor de Língua Inglesa na produção oral do aluno**. 2003. Dissertação (Mestrado em Linguística Aplicada) – Universidade Católica de Pelotas, Pelotas, 2003.

ROSCOE, R. D.; MCNAMARA, D. S. Writing Pal: Feasibility of an Intelligent Writing Strategy Tutor in the High School Classroom. **Journal of Educational Psychology**,

p. 2-16, 2013. Retrieved from http://129.219.222.66/Publish/pdf/Roscoe_2013.pdf on Apr 26, 2018.

RUDNER, L.; GAGNE, P. An Overview of Three Approaches to Scoring Written Essays by Computer. **Practical Assessment, Research & Evaluation**, v. 7, n. 26, p. 1-4, 2001. Retrieved from http://pareonline.net/getvn.asp?v=7&n=26 on Apr 22, 2018.

RUDNER, L. M.; LIANG, T. Automated essay scoring using Bayes' theorem. **The Journal of Technology, Learning, and Assessment**, v. 1, n. 2, p. 4-21, 2002. Retrieved from http://www.jtla.org on Apr 22, 2018.

SCHWARTZ, F.; WHITE, K. Making sense of it all: Giving and getting online course feedback. In: WHITE K.; WEIGHT, B. (Eds.). **The online teaching guide**: A handbook of attitudes, strategies, and techniques for the virtual classroom. Boston: Allyn and Bacon, 2000, p. 57-72.

SHERMIS, M.; BARRERA, F. Exit assessments: Evaluating writing ability through Automated Essay Scoring, **Proceedings of the Annual Meeting of the American Educational Research Association (AERA),** New Orleans, LA, p. 1-32, 2002. Retrieved from https://www.govinfo.gov/content/pkg/ERIC-ED464950/pdf/ERIC-ED464950.pdf on May 10, 2017.

SHERMIS, M. D.; BURSTEIN, J. **Automated Essay Scoring**: A cross-disciplinary perspective. Mahwah, NJ: Lawrence Erlbaum Associates, 2003. 256 p.

SHERMIS, M. D.; RAYMAT, M. V.; BARRERA, F. Assessing writing through the curriculum with Automated Essay Scoring, **Proceedings of the Annual Meeting of the American Educational Research Association (AERA)**, p. 1-31, Chicago, Illinois, 2003. Retrieved from https://files.eric.ed.gov/fulltext/ED477929.pdf on May 9, 2017,

SHUTE, V. Focus on formative feedback. **ETS Research Report RR 07-11**. ETS: Princeton, New Jersey, p. 1-55, 2007. Retrieved from https://www.ets.org/Media/Research/pdf/RR-07-11.pdf on May 9, 2017.

SIEMENS, G. **Connectivism**: A learning theory for the digital age. [Online]. Dec 12, 2004. Retrieved from http://www.elearnspace.org/Articles/connectivism.htm on Jul 14, 2018.

SIMS, R. **Interactivity**: A Forgotten Art? [Online]. Jan 27, 1997. Retrieved from http://intro.base.org/docs/interact/ on Jun 4, 2018.

_____, R. Interactivity for effective educational communication and engagement during technology-based and online learning. In: C. MCBEAT, C.; ATKINSON, R. (Eds.), Planning for Progress, Partnership and Profit. **Proceedings EdTech'98**. Perth: Australian Society for Educational Technology. [Online]. 1998. Retrieved from http://www.ascilite.org/conferences/aset-archives/confs/edtech98/pubs/articles/sims1.html on Jun 14, 2018.

_____, R. Promises of Interactivity: Aligning Learner Perceptions and Expectations with Strategies for Flexible and Online Learning. **Distance Education**, v. 24, n. 1, p. 87-103, 2003.

SKINNER, B. F. **Verbal behavior**. Cambridge: Prentice Hall, 1992.

STREETER, L.; PSOTKA, J.; LAHAM, D.; MACCUISH, D. (2004). The credible grading machine: Essay scoring in the DOD. Retrieved from https://assets.pearsonschool.com/asset_mgr/legacy/200728/IEAResearch-Streeter2002-PKT_665_1.pdf on Apr 7, 2018.

STORCH, N.; WIGGLESWORTH, G. Learners' Processing, Uptake, and Retention of Corrective Feedback on Writing. **Studies in Second Language Acquisition**, v. 32, n. 02, p. 303-334, 2010.

TARONE, E. Conscious communication strategies in interlanguage: A progress report. In: BROWN, H. D. YORIO; C. A.; CRYMES, R. C. (Eds.). **On TESOL '77**. Washington: TESOL, p. 194-203, 1977.

TORI, R. A presença das tecnologias interativas na educação. **Revista de Computação e Tecnologia da PUC-SP**, v. 2, n. 1, p.4-16, 2010.

TRUSCOTT, J. The case against grammar correction in L2 writing classes. **Language Learning**, v. 46, p. 327-69, 1996.

TUCKMAN, B. **Manual de Investigação em Educação**. Lisboa: Fundação Calouste Gulbenkian, 2005.

VANPATTEN, B. The ACTFL proficiency guidelines: Implications for grammatical accuracy in the classroom? **Studies in Second Language Acquisition**, v.8, p. 56-67, 1986.

VANTAGE LEARNING. (s.d.). Retrieved from http://www.vantagelearning.com on Apr 14, 2018.

VANTAGE LEARNING. **MY Access! Efficacy Report**. Newtown, PA: Vantage Learning, 2007. Retrieved from http://www.vantagelearning.com/school/research/myaccess.html on Apr 22, 2018.

VETROMILE-CASTRO, R. A usabilidade e a elaboração de materiais para o ensino de inglês mediado por computador. **Revista Brasileira de Linguística Aplicada**. P. 9-23, 2003.

VYGOTSKY, L. S. **Pensamento e Linguagem**. (e-book). Edição eletrônica: Ridendo Castigat Mores, (s.d.). Retrieved from http://www.ebooksbrasil.org/adobeebook/vigo.pdf on May 12, 2018.

WAGNER, E. D. Interactivity: From Agents to Outcomes. **New Directions for Teaching and Learning**, n. 71, Jossey-Bass Publishers, p. 19-26, 1997.

WARE, P.; WARSCHAUER, M. Electronic Feedback and Second Language Writing. In: K. Hyland and F. Hyland (Eds.) **Feedback and Second Language Writing**. Cambridge: Cambridge University Press, 2006, p. 1-29. Retrieved from http://education.uci.edu/uploads/7/2/7/6/72769947/feedback.pdf on Mar 4, 2017.

WARSCHAUER, M.; WARE P. Automated writing evaluation: defining the classroom research agenda, **Language Teaching Research***,* v. 10, n. 2, p. 1-24, 2006.
WARSCHAUER, M.; GRIMES, D. Automated Writing Assessment in the Classroom. **Pedagogies**, v. 3, n. 1, p. 22-36, 2008. Retrieved from https://www.researchgate.net/publication/254308385_Automated_Writing_Assessment_in_the_Classroom on Nov 7, 2019.

WATSON. J. **Behaviorism**. Chicago. The University of Chicago Press, 1930.

WIENER, N. **The Human Use of Human Beings**: Cybernetics and Society, 2d ed. Garden City, New York: Doubleday, 1954.

WEIZENBAUM, J. **Computer power and human reason:** from judgment to calculation. New York: W. H. Freeman and Co., 1976.

WHITE, C. **Language Learning in Distance Education**. Cambridge: Cambridge University Press, 2003.

YANG, N. D. Using MyAccess in EFL writing. **The proceedings of 2004 International Conference and Workshop on TEFL & Applied Linguistics**. Taipei, Taiwan: Ming Chuan University, p. 550-564, 2004.

YUAN, Z; BRISCOE, T. Grammatical error correction using neural machine translation. **Proceedings of NAACL-HLT**. San Diego, California, p. 380-386, 2016.

Appendices

Appendix A

UNIVERSIDADE FEDERAL DE PELOTAS
CENTRO DE LETRAS E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS - DOUTORADO
PESQUISADORA: PROF. MS. GISELE MEDINA NUNES
ORIENTADOR DA PESQUISA: PROF. DR. VILSON JOSÉ LEFFA

Termo de Consentimento Livre e Esclarecido

As informações contidas neste termo de consentimento livre e esclarecido foram fornecidas pelos pesquisadores Prof. Dr. Vilson José Leffa e Prof.ª Ms. Gisele Medina Nunes com o objetivo de obter a autorização, por escrito, do participante que fará parte de um estudo sobre o desenvolvimento da habilidade escrita em língua inglesa por meio de um programa de avaliação automática da escrita em língua estrangeira. O participante terá conhecimento do que será realizado no projeto e dará sua autorização por livre vontade.

Título do Estudo: Avaliação automática da escrita: Uma perspectiva pós-humana no desenvolvimento da habilidade escrita em LE.

Justificativa: No que tange ao ensino de línguas, tem-se observado o surgimento crescente de ferramentas disponíveis *online* com a finalidade de auxiliar na aquisição de habilidades em língua estrangeira (LE), fomentando um aprendizado mais autônomo e automatizado. Com respeito à habilidade escrita, programas vem sendo desenvolvidos com o intuito de fornecer feedback imediato e, por consequência, automático, para produções escritas, na tentativa de minimizar a carga de trabalho que sua correção demanda. As tecnologias da informação e comunicação – as TIC – aliadas à inteligência artificial tem possibilitado a criação de programas com um feedback mais detalhado e abrangente, incluindo não só a revisão dos textos como também o provimento de feedback formativo, aquele que orienta e instrui o aprendiz, em uma perspectiva de aprendizagem construtivista. No entanto, esse cenário instiga o questionamento do potencial interativo de ferramentas virtuais automatizadas para a aprendizagem de uma LE tendo em vista as limitações que se observam na sua programação em apreender nuances da língua e intencionalidade do aprendiz, a fim de fornecer um feedback que de fato preencha as lacunas necessárias para a aprendizagem de forma mais individualizada em um ambiente cada vez mais marcado pela comunicação em massa e generalizada. Ainda, uma vez que o programa dispõe de uma versão que disponibiliza a formação de grupos de aprendizes como em uma turma e fornecimento de feedback proveniente de um professor mediado pelo programa, será observado como essa dinâmica pode ser desenvolvida em prol de um bom aproveitamento do aprendiz.

Objetivos: 1) Verificar como um sistema automatizado online desenvolvido para a avaliação da escrita em LE se constitui como um recurso com a finalidade de aprimorar a habilidade escrita de seus usuários de forma autônoma baseado em uma relação interativa homem-máquina; 2) Descrever o sistema observando o seu nível de interatividade com o usuário e a influência desse aspecto na aprendizagem; 3)

Examinar o feedback que o programa fornece ao usuário, levando em consideração o público alvo, o tempo de entrega do feedback, os tipos de erro que o programa mais identifica, os tipos de erro que o programa mais corrige e a forma de apresentação do feedback ao usuário; 4) Observar efeitos da dinâmica de feedback automatizado em conjunto com feedback escrito proveniente de um professor no aprimoramento da escrita em LE do aprendiz.

Procedimentos: O participante será instruído e produzirá textos escritos para tarefas selecionadas pela pesquisadora em um programa online específico para produção escrita em Língua Inglesa. Serão lançadas dez tarefas em uma frequência semanal e requerem entre 30 e 200 palavras. Os tópicos, sua complexidade e extensão serão determinados de acordo com o nível de proficiência do participante, o qual será previamente estabelecido. Após, se procederá à aplicação de um questionário contendo perguntas a respeito de como o participante percebeu a correção automatizada combinada com o feedback fornecido pelo professor por meio do próprio programa, além de aspectos do programa que considerou positivos e negativos enquanto o utilizava para produzir seus textos. O questionário será composto de duas partes: a primeira com oito questões fechadas (respostas controladas) sendo necessária somente a marcação do número um ao cinco de acordo com a legenda referente a cada número e a segunda parte abarcará cinco questões abertas (respostas dissertativas).

Desconfortos e riscos esperados: As atividades não apresentarão risco ao participante.

Informações Adicionais: Não haverá identificação do nome do participante, sendo os dados utilizados única e exclusivamente em eventos científicos da área ou áreas afins. É permitido ao participante desistir da pesquisa em qualquer momento. Somente a pesquisadora e o participante tem acesso às produções textuais. Além disso, o participante poderá receber, sempre que solicitadas, informações atualizadas sobre todos os procedimentos objetivos e resultados do estudo realizado. Não haverá despesas financeiras decorrentes da participação na pesquisa.

Eu,
_____,
portador(a) da carteira de identidade nº _____, certifico que após a leitura deste documento e outras explicações fornecidas pelos professores Vilson José Leffa e Gisele Medina Nunes (53) 981284345, sobre os itens acima, estou de acordo com a realização deste estudo e autorizo a minha participação.

_____
Participante

_____ _____
Prof. Dr. Vilson José Leffa Prof.ª Ms. Gisele Medina Nunes