

# English Correction Software

## ToBIT Paper

Tobias Koller

University of Applied Science Northwestern Switzerland (FHNW),  
4000 Basel, Switzerland

**Abstract.** This paper forms ....

**Keywords:** Grammarly · Natural Language Processing · Language Correction Software

## 1 Introduction

Numerous software solutions on the market promise to help in particular non-native English speakers with grammatical error detection and improving the style and structure of their writing. In the course of the module “Innovative Topics in Business Information Technology” (ToBIT), I am going to evaluate different products regarding their functions, ease of use and effectiveness in supporting the writing process. The main goal is to use literature review to discover how effective the correction software is under real conditions.

One widely known and used writing tool for grammar checking, spell checking, and plagiarism is Grammarly®. Since it is one of the leading products in this field and there are innumerable research papers to support, I focus on this particular tool.

After the evaluation of the tools, I will also describe different natural language processing (NLP) techniques that are being applied. In the final part of the document, the findings will be discussed to show clearly the current state of art in this field. A recommendation to University of Applied Sciences and Arts Northwestern Switzerland (FHNW) will be given on what products to consider or how to develop a new product internally.

## 2 Introduction

The following chapter

### 3 Existing solutions

#### 3.1 Grammarly

Grammarly, a company with its eponymous popular language correction software, aims to improve the communication of people. Max Lytvyn, Alex Shevchenko, and Dmytro Lider founded the company 2009 with a strong focus on supporting the student's writing process. In the meantime, they broadened their scope to businesses as well as personal writers. With Grammarly @edu and Grammarly business, they have their dedicated divisions for those clients. Their correction service is based on analysing the written content in real-time while showing suggestions in the form of correction cards. With a simple click, the changes can be applied or rejected. To promote learning of the writer further grammatical information regarding the specific issue can be retrieved from the card which can help to decide on whether to accept the changes or not. ("Write your best with Grammarly." n.d.)

Being a successful grammar checker requires to be available where people create texts. With such a diverse client group, those places are multi-faceted. While academic and business end-users might use classical text processing software, most personal use cases are not so easy to unify. Web forums, chat messages, e-mails or social-media platforms are some of the areas users compose content. Grammarly responds to this with various products. Browser plugins for all major browsers (Chrome, Safari, Firefox, Edge) provide accompanying grammar service during your browsing experience. Input text fields are automatically detected and errors are highlighted with the option to see the correction card. Users of Windows and Mac can download a Grammarly text editor, but similar functionality can also be used in their web application on [app.grammarly.com](http://app.grammarly.com). To also serve customers who write on mobile devices, the Grammarly team developed a keyboard for Android and iOS. For both Microsoft Word and Outlook, Grammarly provides a special plugin that integrates directly and seamlessly with the writing process.

Which type of help can writers expect from the software? According to their webpage ("Write your best with Grammarly." n.d.) support ranges from grammar checking, tone detector and for paying users even a plagiarism checker. They claim that their grammar checker does not only find misspelt words but also recognises comma and other punctuation misuses. Furthermore, their premium service includes more advanced suggestions to enhance the writing style. To help writing an appropriate language one can define so-called goals on which the algorithm bases its recommendations on. Parameters that can be predefined include target-audience, formality, domain and tone of writing.

Users can immediately start using Grammarly with just a simple registration. The usage of the base functionality is available for free for an unlimited time

span. This lowers the barrier of entry immensely, since all a user need to have is an e-mail address and a computer or smartphone with an internet connection. However, a premium service is offered against payment. Grammarly’s website (“Write your best with Grammarly.” n.d.) shows the deviation from the premium plans to the free version. The plans “Premium” and “Business” both comprise the same advanced features including suggestions in the following categories.

- Fluency
- Readability
- Engagement
  - Compelling vocabulary
  - Lively sentence variety
- Delivery
  - Confident language
  - Politeness
  - Formality level
  - Inclusive language
- Plagiarism detection

“Business” plan only differs in the account management tools that is available to the organisation’s administrator as well as business-oriented billing. The available software and plugins (browser, MS Word, MS Office, mobile keyboard) are equally available for paid and free users. Since Grammarly has its focus on supporting students they target educational institutions with their programm “grammarly@edu”. The grammar checking functionalities seem to be the same as in the “Premium” and “Business” plans but they offer specialised licenses and 24/7 support. Different versions of the same tool make the evaluation of the tool in the following chapter more difficult since some of the research found is based on the limited functionalities. Furthermore, the software developed greatly in the past years as can be seen on printscreens from Dembsey (2017). This need to be taken into account when judging the results from this research.

*some notes to the author*

## 4 Evaluation

### 4.1 Grammarly

The Grammarly’s web presentation (“Write your best with Grammarly.” n.d.) makes strong claims about the effectiveness and usefulness of the system. Over the past years several researcher attempted to measure the impact of using

Grammarly on the student's written performances and determine the overall quality of the feedback produced. Others (Cavaleri & Dianati, 2016) aimed to determine the perceived ease of use and perceived usefulness to answer the question if this technology will be accepted or not. The methods used include comparing Grammarly's feedback to the one provided by online writing consultants (Dembsey, 2017), comparing student's performance in language tests before and after being exposed to the software (Qassemzadeh & Soleimani, 2016) and different kinds of surveys and questionnaires (Nova, 2018) (Cavaleri & Dianati, 2016) (Ventayen & Orlanda-Ventayen, 2018).

**Quality of feedback** Feedback provided by a grammar correction software should be understandable by the end user in order to promote learning. Simply accepting suggestions blindly without questioning support in avoiding the same error in the future. Additionally, the correction algorithms are not flawless and their proposed changes should always be questioned. This is only possible if the writer understands the issue that was found in his writing (Dembsey, 2017).

The understanding of the problem and thereof the possible learning gain is dependent on whether the student understands the terminology used in the feedback. In grammar situations and constellations often can be described very precisely by a specialised term. Especially for students in English as a foreign language (EFL), those terms might be ambiguous and need further explanation. Dembsey (2017) found that Grammarly used 52 different terms in the correction of three essays while on the other hand 10 online writing consultants used only 32 terms (all consultants combined), or 10 terms on average, for the same documents. Furthermore the consultants used much more accessible language in their comments' explanation and even attempted to use the student's language to give more comprehensible feedback. Giving feedback in an appropriate format for the receiver can be achieved by humans way better than by algorithms. In general advanced terminology is not supportive for the learning process of the student. Simple language should be used whenever possible (Dembsey, 2017).

In the best case the the grammar correction software's feedback encourages the user to scrutinise the passage of attention and give some valuable recommendation to improve it. However, when not detecting the issues correctly, misleading feedback can lead to the author's confusion. In an interview conducted with Indonesian EFL postgraduate students (Nova, 2018) multiple participants reported that Grammarly changed the sentence's intentional meaning and therefore led to their confusion. As long as students are aware of the software's mistake they can simply ignore it and proceed. More harmful are those incidents when the student is at a beginner-level in English. He might be more tempted to accept any changes proposed without the lack of experience spotting those erroneous suggestions. As Vojak, Kline, Cope, McCarthey, and Kalantzis (2011) (**!!! isn't that a bit too long of an author's list??!!!**) points out such a situation of uncertainty is counterproductive to the author's development of confidence in

his writing. The notion of something being wrong with his writing motivates to comply with general phrases and standard structure in the future. Instead of promoting better writing style, they “fear that the persistent underlying urge towards conformity may stifle individual creativity” (Vojak et al., 2011).

Grammar correction on a sentence level follows rather clear rules whereas connections within a paragraph or even the logical structure of the whole document are much more sophisticated tasks. Unsurprising that also automated correction software like Grammarly have their difficulties. Students experienced the lack of context aware checking like coherency and cohesiveness within a text. Those who needed the software only to check the grammar did not find this an issue (Nova, 2018). The same results were found by Dembsey (2017) who observes that Grammarly treats each word and sentence individually and not making any connections between them, therefore drastically reducing the learning opportunity compared to expert feedback.

**Amount of errors found** On first sight high quantitative figure of detected errors seems to demonstrate the quality of a correction algorithm. We will see why this appearance might be deceptive.

During the comparison of Grammarly and writing consultants in analysing three student essays Dembsey (2017) Grammarly observed a total of 118 issues whereas the cumulative average of the 10 consultants only brought up 51. Repetition of the same issues was the main driver for such a high number of detected issues. A human proofreader could encourage the student to look for additional instances of the same mistake by themselves, leaving more time for different issues. In order to get a better view of the issues discovered, all issues were categorised which led to a total of 16 categories to which every issue could be assigned. In all the essays combined Grammarly’s correction cards could be assigned to only six types of issues. This again show the rather narrow range of recommendations. Cumulated all 10 consultants addressed 15 issue categories and even on average they addressed more (8) diverse topics than Grammarly.

Despite having found more issues than human proofreader, Grammarly’s issue detection was highly repetitive and only addressed a narrow range of issues. The consultants used less comments but gave more in-depth explanation and could even connect sentence level issues to general (thesis) level issues. Furthermore, a high number of issues is often not beneficial for the learning rate of students, as they might become intimidating and demotivating. (Dembsey, 2017)

**Accuracy** A more crucial measure of value provided by feedback than number of issues detected is the accuracy of the results. False positives are reported issues that are no problems at all. Dembsey (2017) also considered incorrect use of term or incorrect explanation as inaccuracy. 41% of Grammarly’s correction

cards where inaccurate, either being false positives or using wrong terms for the specific issue. At the same time consultants only had an average inaccuracy of 10% which originated mostly from using wrong terms. (Dembsey, 2017)

The decision if an issue should be raised or not is also dependent on the type of writing. This puts an automated correction software in a disadvantageous position, since detecting type of writing as well as target audience is generally difficult. Cavaleri and Dianati (2016) tested Grammarly's premium version and could indicate the type of writing. For "essay", "dissertation", "presentation", "blog", "business document" or "creative writing" different rules of raising issues would be applied. This improved the accuracy of the feedback profoundly.

At the moment of writing Grammarly also allows setting some meta information to the document allowing for increased accuracy. Audience (general, knowledgeable, expert), formality (informal, neutral, formal), tone (neutral, confident, joyful, optimistic, friendly, urgent, analytical, respectful) and intent (inform, describe, convince, tell a story) are available in the free version. The latter two are marked as experimental. Only the domain (academic, business, general, technical, causal, creative) is only available in the premium version. Seeing those features, specially the one being experimental, shows that Grammarly has already detected the necessity to increase accuracy by means of better contextual issue detection.

### **Perceived Ease of use**

- ease of access as well (Nova, 2018) page 7
- low barriers, fast and quick setup
- extremely easy to use (Cavaleri & Dianati, 2016) page 12
- terminology used might be a barrier for some, requires a lot of conceptual thinking (Cavaleri & Dianati, 2016) page 12
- SUS System Usability of Scale (Ventayen & Orlanda-Ventayen, 2018)

### **Perceived Usefulness** Usefulness Feedback (Nova, 2018) page 6

- helped learning from mistakes
- examples helped to understand the issue better
- rises awareness of errors
- more advanced writer benefit most (Cavaleri & Dianati, 2016) page 12
- repeatedly mentioned that it also checks the references which is useless since one cannot simply make changes (Ventayen & Orlanda-Ventayen, 2018) page 19 and (Nova, 2018) page 8
- specially useful to non native speakers (Ventayen & Orlanda-Ventayen, 2018) p22

- supplement rather than substitute (Ventayen & Orlanda-Ventayen, 2018) p22

## 4.2 Application in classrooms

Li, Link, and Hegelheimer, 2015

- instructors view and usage of AWE
- students view highly depend on the instructors approach

## 5 Techniques of natural language processing

## 6 Discussion and recommendation

## 7 Self experiment

Just cite Dembsey, 2017

Text cite Dembsey (2017)  
paren cite (Dembsey, 2017)

citep paragraph? And so on (Dembsey, 2017) The following chapter  
text quote “word for word ” (Dembsey, 2017)

## 8 Bibliography used

The list of references is not final and will be extended during the process of writing the ToBIT paper.

## References

- Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, 10(1), A223–A236.

- Dembsey, J. M. (2017). Closing the Grammarly® Gaps: A Study of Claims and Feedback from an Online Grammar Program. *The Writing Center Journal*, 36(1), 63–100. Retrieved October 14, 2019, from <https://www.jstor.org/stable/44252638>
- Li, J., Link, S., & Hegelheimer, V. (2015). Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction. *Journal of Second Language Writing*, 27, 1–18. doi:10.1016/j.jslw.2014.10.004
- Nova, M. (2018). UTILIZING GRAMMARLY IN EVALUATING ACADEMIC WRITING: A NARRATIVE RESEARCH ON EFL STUDENTS' EXPERIENCE. *Premise: Journal of English Education*, 7(1), 80–97. doi:10.24127/pj.v7i1.1332
- Qassemzadeh, A., & Soleimani, H. (2016). The Impact of Feedback Provision by Grammarly Software and Teachers on Learning Passive Structures by Iranian EFL Learners. *Theory and Practice in Language Studies*, 6(9), 1884–1894. doi:10.17507/tpls.0609.23
- Ventayen, R. J. M., & Orlanda-Ventayen, C. C. (2018). *Graduate Students' Perspective on the Usability of Grammarly® in One ASEAN State University* (SSRN Scholarly Paper No. ID 3310702). Social Science Research Network. Rochester, NY. Retrieved October 23, 2019, from <https://papers.ssrn.com/abstract=3310702>
- Vojak, C., Kline, S., Cope, B., McCarthey, S., & Kalantzis, M. (2011). New Spaces and Old Places: An Analysis of Writing Assessment Software. *Computers and Composition*, 28(2), 97–111. doi:10.1016/j.compcom.2011.04.004
- Write your best with Grammarly. (n.d.). Retrieved November 27, 2019, from <https://www.grammarly.com/>