

Technical Disclosure Commons

Defensive Publications Series

September 30, 2019

User-Tailored Language Training

Marcos Calvo Lance

Joan Pastor

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Calvo Lance, Marcos and Pastor, Joan, "User-Tailored Language Training", Technical Disclosure Commons, (September 30, 2019)
https://www.tdcommons.org/dpubs_series/2531



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

User-Tailored Language Training

Overview

Generally, the present disclosure is directed to systems and methods for language training. Language training can be tailored to a user's skill level in the language. In particular, in some implementations, the systems and methods of the present disclosure can include or otherwise leverage one or more machine-learned models to learn a user's skill level in a language based on user communications in the language. Additionally, the one or more machine-learned models can provide tailored exercises to a user based on one or more learning objectives associated with the user.

Introduction

Mastering a language can be a desirable skill but is often wrought with challenges and difficulties. For instance, languages can be nuanced and have many aspects that require a deep understanding of other aspects of the language. By way of example, many concepts in a language cannot be fully understood or can otherwise be confusing without knowledge of grammar rules, vocabulary, expressions or idioms, practical usage, or other concepts. Existing technologies to assist with language training typically teach based on a curriculum of language. While a user of these existing technologies can obtain an understanding of languages, the curriculum can fail to provide personalized feedback to the user to allow the user to focus on areas of particular interest or difficulty to the user and, thus, obtain a deeper understanding of the language.

Summary

Generally, the present disclosure is directed to systems and methods for user language training. The training can be tailored to the user's skill level in the language. In particular, in some implementations, the systems and methods of the present disclosure can include or

otherwise leverage one or more machine-learned models to learn a user's skill level in a language based on user communications in the language. Additionally, the one or more machine-learned models can provide tailored exercises to a user based on one or more learning objectives associated with the user.

Detailed Description

According to the present disclosure, a user can be provided with a personalized language training method that can be adapted to a language skill of the user. For example, one example implementation of a personalized language training method in accordance with the present disclosure is illustrated in Figure 1.

In some implementations, the personalized language training method can include establishing a training language and/or an initial user skill in the training language. For instance, a user can select the training language as a language for which the user wishes to gain instruction. An initial user skill level in the training language can be established according to any suitable method. For instance, in some implementations, the user can indicate an estimated initial user skill level in the language. For instance, the user can select an initial user skill level from a discrete list (e.g. "Beginner," "Intermediate," "Expert") or select a numerical value from a range (e.g. from 1-10). In some implementations, a level test can be provided to the user to determine the initial user skill level in the training language. For instance, the level test can include one or more questions, exercises, or other suitable assessment metrics to score the competency of the user in communicating in the training language. In some implementations, a machine-learned model can be configured to predict the initial user skill level based on recent communications from the user (in that language). For instance, the machine-learned model can be the user skill

model as described below. In some implementations, the user can choose to allow or restrict the use of some or all communications to determine the initial user skill level in the language.

According to example aspects of the present disclosure, a computing system can obtain one or more communications from a user. For instance, the one or more communications can be provided from the user to a user device and, with consent of the user, can additionally be provided to the computing system, which may be or otherwise include a part of the user device, such as another application on the same user device, another user device, a remote computing device, or any other suitable computing system.

The one or more communications can have an associated language. For instance, the one or more communications can be obtained after a training language and initial user skill in the training language have been established. Thus, one or more communications can be obtained that are in the training language. Any suitable communications can be used in accordance with the present disclosure. For example, the one or more communications can include textual communications such as, but not limited to, emails, social media posts, texts, chatroom messages, or other suitable textual communications. As another example, the one or more communications can include audial communications such as, but not limited to, phone calls, audio messages, speech recognition data, voice commands, or other suitable audial communications. In some implementations, the user can choose to allow or restrict some or all types of communications from being obtained. For instance, the user can choose to only allow textual communications (or a subset thereof) to be obtained. As another example, the user can choose to restrict communications to or from certain addressees.

In some implementations, the one or more communications can be provided by the user directly to the computing system. For instance, the user can select one or more communications

that he or she wishes to provide to the computing system. In some implementations, the one or more communications can be obtained automatically by the computing system, with consent from the user. For example, the computing system can obtain one or more communications from sources matching some user-specified criteria.

In some cases, the one or more communications can have an associated formality. For instance, formality can be a relative measure of the tone, setting, etc. of the communication. For example, a phone conversation between friends could have relatively low formality while a business email could have relatively high formality. In some implementations, the computing system can identify the formality associated with some or all of the one or more communications.

According to example aspects of the present disclosure, a computing system can identify one or more learning objectives associated with the user. The one or more learning objectives can be identified to, for instance, improve the user's understanding of and/or competency in the training language. For instance, the one or more learning objectives can be identified based on one or more error patterns in the one or more communications. For example, the one or more error patterns can be indicative of one or more recurring errors in the one or more communications from the user, such as errors occurring across multiple communications and/or from multiple different sources of communications. The errors can be identified by any suitable method(s) in accordance with the present disclosure. For example, the errors can be identified by parsing the one or more communications (e.g. by grammar heuristics and/or by neural networks). As another example, the errors can be identified by spell checking algorithms. As another example, the errors can be identified by language models (e.g. by n-grams and/or by neural networks). As another example, the errors can be identified by dictionary comparison and/or by

semantic disambiguation techniques, such as by dictionary-based semantic disambiguation techniques.

In some implementations, the one or more learning objectives can be associated with grammatical or other linguistic aspects that are not necessarily errors. For instance, one example learning objective can include an improved vocabulary, such as learning a new word. As another example, a learning objective can include developing an understanding of idioms. As another example, a learning objective can include developing stylistic components of language, such as sentence structure, length, organization, or other suitable stylistic components. For example, the learning objective can be selected to teach a user to diversify repeated words, phrases, structural elements, etc., in the user's communications. As another example, a learning objective can include developing an understanding of grammatical rules for appropriate sentence structure/syntax.

Each type of learning objective can have an associated difficulty or required skill. For instance, the associated difficulty or required skill can be related to how difficult it is for a person to grasp the concept associated with the learning objective. It can be desirable to avoid learning objectives associated with concepts that are difficult for the user to grasp at the user's current skill level. For instance, training exercises that are too difficult for the user to complete can serve to confuse or demoralize the user. As such, even if a learning objective is identified in the user's communications, it can be desirable to postpone training the user based on the learning objective or otherwise ignore the learning objective such that training exercises and/or progress reports based on the learning objective are not provided to the user.

Thus, the one or more learning objectives can additionally and/or alternatively be selected based on a current user skill level in the training language. For instance, systems and

methods according to the present disclosure can employ a skill model configured to learn the current user skill level in the training language. For instance, the skill model can be a machine-learned model. For instance, the skill model can receive the one or more communications as input and predict the current user skill level as output.

As another example, the skill model can receive one or more candidate learning objectives and/or a current user skill level as input and provide one or more selected learning objectives as output. For instance, the one or more candidate learning objectives can be a set of some or all learning objectives identified from one or more communications from a user, as discussed above. The one or more selected learning objectives can be a subset of the one or more candidate learning objectives. For instance, the one or more selected learning objectives can be selected based on the current user skill level such that the one or more selected learning objectives are tailored to facilitate the user understanding the one or more selected learning objectives.

For example, one example learning objective could be “proper use of past tense.” A second example learning objective could be “use of passive voice.” It can be determined that, for example, “use of passive voice” has a higher required skill level than “proper use of past tense.” Thus, a user at a relatively low skill level might provide a communication having errors in both past tense use and passive voice use, but would only be able to realistically improve upon past tense use based on the user’s current skill level. As such, the skill model could predict the “proper use of past tense” learning objective as a selected learning objective for this user.

In some implementations, the formality of the one or more communications can be considered in determining the one or more learning objectives. For example, communications having low formality may be exempted from the determination of learning objectives due to

being, for example, hastily spoken or written, or otherwise in a context where the user is unlikely to consider grammatical correctness important.

According to example aspects of the present disclosure, a computing system can provide one or more exercises to the user based on the one or more learning objectives. For instance, the one or more exercises can be provided based on the selected learning objectives as described above. The one or more exercises can provide instruction related to the one or more learning objectives. For example, the one or more exercises can relate to correcting an error that is commonly present in the user's communications. As another example, the one or more exercises can serve to illustrate a linguistic concept, such as grammar rules, vocabulary, stylistic concepts, or any other suitable concept.

In some implementations, a machine-learned model (e.g., a generative model) can be configured to generate the one or more exercises. For example, the machine-learned model can be trained on existing exercises and extrapolate new exercises from the existing exercises. The model can also be trained on a sample of the user's communications, to tailor the training exercises. Example exercises can include, but are not limited to, quiz questions (e.g., multiple-choice questions), cloze tests, pronunciation tests, or other suitable exercises. In some implementations, the one or more exercises can be tailored based on formality. For example, the one or more exercises can serve to indicate appropriate stylistic conventions for different formality settings.

As the user continually interacts with the system, the user skill can be continually learned and updated. For instance, the one or more communications, one or more learning objectives, one or more exercises, and/or any data derived from the preceding, can be provided to the skill model at any of the steps described above to continually update the model. In this way, the

system can track progress of the user. For instance, the skill model can learn that an error pattern that was present in earlier communications has been corrected in later communications, indicating that the user has satisfied the learning objective. Moreover, as explained above, this technology can unlock further learning objectives that may have previously been considered too advanced.

In some implementations, the computing system can provide a summary to the user. The summary can be provided by any suitable method, such as, but not limited to, email, notification (e.g. push notification), text message, audial message, data field entry, or any other suitable method. The summary can be provided at any suitable point in time. For instance, the summary can be provided in response to an action of the user, such as the user starting an application, or by user request. As another example, the summary can be provided in a recurring manner, such as daily, weekly, biweekly, monthly, bimonthly, yearly, or in any suitable recurring manner. In some implementations, the summary can be provided along with some or all of the one or more exercises. In some implementation, different data and/or types of data can be provided with each summary.

In some implementations, the summary can include data related to a user's progress in learning the training language. For example, the summary can include a list of some or all of the learning objectives that have been completed and/or are selected for training. For example, the summary can list all learning objectives completed in some past time duration, such as in the past week or month, and/or in the time since the last summary was provided. In some implementations, the summary can include selected excerpts from the one or more communications. For instance, the selected excerpts can be selected to illustrate errors in the user's communications. For example, if the summary is provided along with an exercise, the

summary can include communications with errors relevant to the learning objective associated with the exercise. In some implementations, the user can restrict these snippets to some specific type of communication or addresses.

Systems and methods according to the present disclosure can be implemented on any suitable user device. Example user devices include, but are not limited to, desktop computers, laptop computers, tablet computers, phones, smart speakers, web applications, or any other suitable user device. In some implementations, the systems and methods according to the present disclosure can span more than one user device. For instance, in one implementation a system having a web app accessible via a computer can be employed in combination with installable applications on, for example, smartphones and/or smart speakers. In some implementations, different types of communications can be obtained at different devices. For example, aural communications might be obtained from a smart speaker. Additionally and/or alternatively, textual communications might be obtained from an application on a computer and/or smartphone. In some implementations, the user can choose to allow or restrict collection of communications from some or all user devices.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs, or features described herein may enable collection of user information (e.g., information about a user's textual and/or aural communications), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user so that a personal identity of a user cannot be determined. Thus, the user may have control over what information is collected

about the user, how that information is used, and what information is provided to the user. Additionally, the user can be provided with controls to enable and/or disable services according to example aspects of the present disclosure at any time.

The technology discussed herein makes reference to servers, databases, software applications, and other computer-based systems, as well as actions taken and information sent to and from such systems. The inherent flexibility of computer-based systems allows for a great variety of possible configurations, combinations, and divisions of tasks and functionality between and among components. For instance, processes discussed herein can be implemented using a single device or component or multiple devices or components working in combination. Databases and applications can be implemented on a single system or distributed across multiple systems. Distributed components can operate sequentially or in parallel.

Thus, while the present subject matter has been described in detail with respect to various specific example implementations, each example is provided by way of explanation, not limitation of the disclosure. Accordingly, the subject disclosure does not preclude inclusion of such modifications, variations and/or additions to the present subject matter as would be readily apparent. For instance, features illustrated or described as part of one implementation can be used with another implementation to yield a still further implementation.

Figure

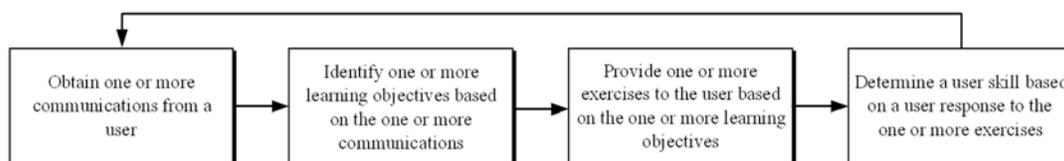


Figure 1