Personalized Learning for practicing an English reading test

Donghee Hong & Adrien Di Pasquale
Department of Knowledge Service Engineering
KAIST – Daejeon, South Korea

December 19, 2012

Contents

1 Motivation

Everyone has their own way of speaking, thinking, reacting. However, classic education systems are "one size fits all" : the same teacher, the same material and the same methodology for a given group of students. As freedoms and technologies evolve and allow knowledge to be less centralized, education is shifting towards new paradigms.

Personalized self-education (or autodidactism) is one solution to the "one size fits all" problem, as it allows a student to be independent and follow his own pace. The amount of resources available on the internet has also contributed to this method's expansion. However, although you are free to choose the material you are going to use among those available, you usually cannot find or generate a truly personalized material.

A key factor for successful education is motivation, and different students are not likely to be motivated by the same things. It is common to say that it is more motivating to learn something that relates to the real life, or to something interesting. It is indeed rewarding to be able to apply what you learn. However, all students have different life styles and activities, different interests.

Our general hypothesis is that students would be more motivated if the material relates to their daily life, their interests, or if the material is formatted in a way they appreciate.

Here are some examples of personalization materials :

1. a classic language lesson based on a dialog of a real life situation with the students' friends' names and pictures instead of fictional characters

2. a signal processing lesson providing examples for a song the student likes, or a photo, a video.

3. a lesson on a very abstract scientific matter mimicking the speech of someone the student likes, or using a type of humour the student appreciates or a idiomatic sentences.

These are three dissociated kinds of material personalization : *real-life*, *interests* and *taste*. These different methods could be more or less appropriate for different students, or different types of lessons. They could also have different impacts on the levels of student's engagement and motivation and on the overall material efficiency (which is correlated).

Our research will try and assert that personalization materials can be more efficient and engaging. We will not try to compare the different methods, but rather check that the common concept is valid.

Our research will focus on teaching English as a foreign language. This is a very appropriate field, since any of the personalization methods can be easily applied. However, we will only explore personalizing the material contents with the students' information and interests. Because of the complexity of such systems, we will not research customizing the materials' formats and style.

2 Background - Related Work

These days online learners can be diverse, from different backgrounds, and with a variety of different learning goals, motivations, learning skills and styles. Even though the online learners are various, most of the current web-based learning systems are still delivering the same educational resources in the same way to learners with different profile [3]. Personalization involves using technology to accommodate these differences. It is becoming an increasingly important factor in education. However, it raises many technical and cognitive challenges. Many approaches have been attempted by researchers in the past decades.

Virtual Learning Environments (VLEs) are web-based systems that give a lot of freedom to the users and allow interaction. Researches have been conducted in this field to integrate personalization of the materials and learning plans. Xu and Wang developed a personalization model, implemented it in a VLE and tested its results [8]. The outcome was that the personalized system yielded significantly higher scores for most of the tests, and that the students appreciated it more than the classical one.

Studies have been conducted to measure the personalization effects, especially on low level students. Lopez and Sullivan investigated the effects of personalized instruction on the achievement of eighth-grade Hispanic students from a rural junior high school on mathematics problems [4]. The objective was to prove that personalizing story problems would make their context more familiar to the students, and therefore improve their performance. The results of this research were significantly positive. It was also found that boys had lower results with standard materials than girls and that personalization increased their results more significantly than for girls.

Adaptive Learning is a field closely related to Personalized Instruction. It makes use of computers to deliver learning content that is personally adapted to the current learner. There are various sources of personalization information that can be used to create such content : the learning style, the cognitive style or learning achievement. In the context of a recent experiment run by Chung-Hua researchers (Taiwan), an Adaptive Learning system was built and

evaluated [6]. Their approach was innovative : they mixed two sources of personalization information, whereas previous systems usually used a single one. The analysis of their results showed that both single-source and double-source systems could improve learning efficacy. They also found out that only the double-source would significantly improve the learning efficiency.

There are also papers covering technical parts of the personalization process. Khribi & Jemni investigated automatic personalization by profiling users on their web navigation history and preferences similarities [3]. Instead of applying this personalization method to build an actual CBT System, they computed personalized lists of recommended links to online materials. To do this they used information retrieval techniques and data mining on both the user profiles and the materials contents. It is however hard to evaluate their results, as it is often with IR systems.

In the field of teaching english as a foreign language, some approaches of personalized instructions were tested. Horie, Kashiwabara, Yamaguchi and Iijima developed a Glossary Generator [2] that takes into account the students level and the materials' difficulties. It also outputs some personalized teaching materials where the words unknown to the current student are highlighted. Its effectiveness has not been measured, though. Closely related is Chen, Hsu, Li, and Peng's research paper : it models a learning system that recommends English news articles depending on the student's reading skills [1]. They used a custom fuzzy algorithm to compute the student's skills based on his answers to short surveys.

These two last studies show interest for personalization in the field of foreign language teaching but the systems can only generate 'simple' learning materials : lists of words or links. Generating a complete personalized material is a big challenge, as raised by Turker, Görgün and Conlan in "The Challenge of Content Creation to facilitate Personalized eLearning Experiences" [7]. In this paper they introduce the European project iClass, a framework to build personalized eLearning systems. It provides services to model the user's profile and preferences and other services to generate personalized contents and learning strategies.

3 Methods

3.1 Hypothesis

"practicing with personalized learning materials compared to non-personalized learning materials will improve students scores on English reading comprehension test, because students will be more engaged with stories related to

themselves"

Vygotsky Learning theory includes the concept of a Zone of Proximal Development. In our experiment, we followed this concept : by transforming some of the complex context parts, we allow users to learn step by step. This approach should therefore yield better results than a traditional approach.

3.2 Experiment

Our test users will be around 30 KAIST students, only native Koreans. They are around 22 - 30 years old, have a strong scientific background and most of them also have a good level of english. We will aim at those with a TOEIC level under 890, because it has been noted that lower level students tend to be more sensitive to personalization.

Our experiment is composed of 5 steps : *registration*, *pre-test*, *practice*, *post-test*, and *questionnaire*. The pre-test will be performed unmonitored at home, about one week before the experiment. The whole process will take place in the lab, under our supervision. For the training step, the students will be divided in two groups : the control group will use standard materials and the study group will use personalized ones (between group test). During this central step, we will measure the time spent, the results and the students' attitudes. None of the 5 steps will have a time limit, however for the pre and post-tests only, we will tell the students that they should complete it quickly, as in a real test. For the pre and post-tests we will use standard materials for both groups. The reason for this is that we want our experiment to focus on the results of personalized vs standard training materials for a real-life test (hence a standard test).

The customized material will be generated for each student based on personal information like the name, age, address, but also on personal interests or point of views : are you more of a family person or a career oriented one ? would you like to live close to the beach or the mountain ? do you have a strong passion for music ? ... Students will fill an online form to let us know this. Only the Study Group data is going to be used, but for equity the other group will take this survey as well. Our standard materials will consist of TOEIC step 7 tasks, slightly incremented with context sentences. These sentences will be as neutral as possible : they should not help the students answer the questions. Then, the personalized materials will be generated by replacing these new context sentences contents (and other pre-existing identified context sentences) with contents based on the user profile.

The task consists of a document of about 100 words, often a letter, but sometimes an ad or another piece of informative paper, and several questions related to it. The questions all have 4 choices, with a single correct one. They

are usually not straightforward questions, so one has to understand the whole context, and also some implicit meanings. This task is copied on TOEIC's part 7.

However the whole process, pre and post-tests training, and tests will be performed on a web-based interface, which removes the classic paper interaction (underlining, circling ..) and might make the task more difficult. We still believe it is better to do it this way, since most of the personalized systems available today are web-based, and also because being able to read fluently on a screen is a useful skill.

4   Results (of the pilot test)

We did not have the time or the resources to conduct the whole experiment as planned. However, we did perform a pilot test with 4 test users matching our criterias. This provided us with valuable feedbacks for the interface, and let us confirm or reject our experimental design.

4.1   Statistical Analysis

We wrote a R script to analyze statistically the results of the experiment. The results we show below come from running this script on our pilot test data, and should therefore *NOT* be interpreted as statistically strong, considering the very low number of test users.

Here is an informative graph of the mean ratio of correct answers for the control (traditional material) and test (personalized material) groups :
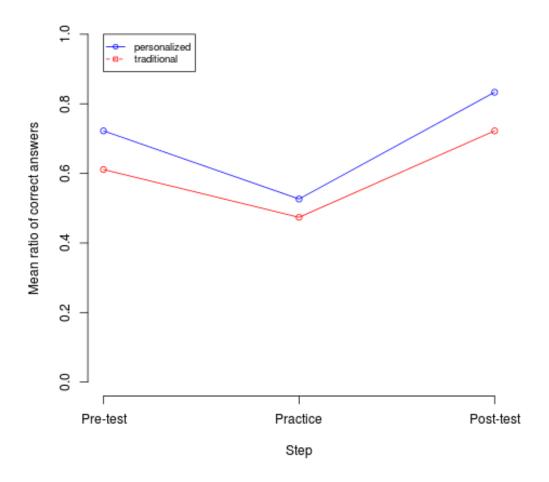
Figure 1: Mean ratio of correct answers

We can see that the personalized material has yielded slightly better results. However this is a biased result, since it doesn't account for the users initial proficiency.

Thus, we computed the ratios of improvement between pre and post-test. The practice results are not used in our computation, since our hypothesis puts the practice only as a mean to achieve better results in the final test. We also computed the ratios of improvement in terms of time spent on the two steps. The results are plotted in figure 2.
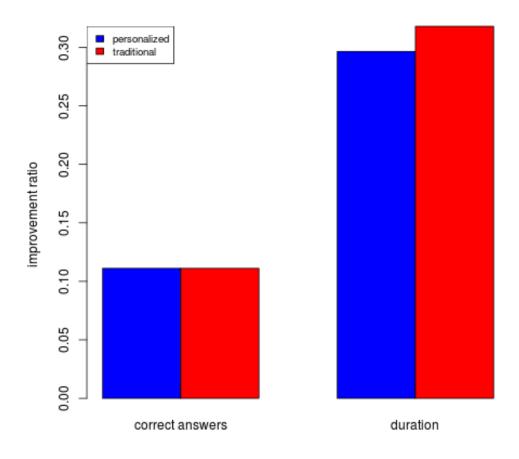
Figure 2: Mean ratio of improvement for time spent and correct answers

Correct answers improved exactly of the same ratio in the control and test groups. This may indicate that the only improvement comes from a difference of difficulty between the pre and post-tests, not from the practice step. Or this could also indicate that the personalization of the practice step had no effect at all on the students results on the post step. Duration improved (i.e. reduced) slightly more for the traditional group. However the big amount of improvement (30%) may indicate that students are just more careful at the beginning of the test and then get bored and conclude quickly in the post test part. Also, some data might be missing from this equation, as

the pre and post test materials have different lengths, and possibly a different number of questions.

To validate our hypothesis, we would have performed a Welch T-Test (different variances) to test the null hypothesis that the correct answer ratio improvement of the control and test groups have the same means. However, there is no need to perform it with our results, since we already know for a fact that they are the same (unfortunately).

4.2 Questionnaire

The questionnaire results were very important for us, especially for this pilot test. Here are the few qualitative results that we can draw from the answers :

1. Personalized group feels significantly more familiar with the contents.

2. Personalized group finds practice material a bit easier and a bit more plausible.

3. Some users believe personalization is fun, but doubt it can really help them.

And some feedbacks for the interface and experiment process :

1. Use radio button instead of checkbox since they are single choice questions

2. The control group is frustrated to receive a 'classic' TOEIC test, they expect personalized content. That is because we made the mistake to let everyone know what the experiment was about.

3. The interface should provide more information on the current position in the experiment.

4. It tends to be stressful because too similar to a real exam

5 Extra Material : Prototype Presentation

5.1 Why build a website ?

Creating the personalized material for each of our students would be very time consuming without the help of a computer. Moreover, we have to make

sure that the Personalized and Traditional materials vary only on very specific parts and in specific ways. Doing it manually could result in a natural tendency to cross these boundaries, and personalize bigger parts of the material. Using an algorithm allows us to be 100% sure of the extent of the personalization.

Our task (TOEIC exercise type 7) is very well suited for a computer use : it does not require any interaction with an instructor, and the submission of a simple form of checkboxes fully characterizes an answer.

It is actually common to perform these tasks on a computer : TOEIC computer-based training softwares are widely available. It is also very probable that the test itself will soon be performed using computers.

A paper-based version of this task would have had a single extra benefit : the ability to highlight and annotate the material, which students often find useful.

5.2 Website walkthrough

5.2.1 At Home

The first step can be done by the user at his/her home. This allows to reduce the in-lab steps complexity and duration.

1. The user receives a mail

   This mail is an invitation to join the experiment. It contains a link to the website and a predefined login. Users should follow the link and sign in using this login. They will then see a welcome message and a single link towards the profile form.

   Some general remarks about the website at this point : it has a minimalist design, and a lot of space. Font sizes are used to emphasize on the important parts. The header color depends on the user's current step.

2. The user fills his / her profile

   It is composed of various fields on different personal information : age, size, hobby, favourite artist ... The fields are either open or provide a fixed set of choices.

   All the users have to go through this step, even those who will be put in the control group. This ensures equity over the two groups : the students should not directly know which group they belong to.

   After filling this form, the website displays a message reminding them of the rendez-vous time and location for the following in-lab steps.
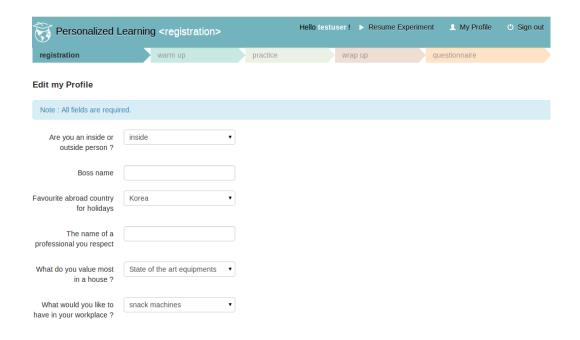
Figure 3: User profile page

5.2.2 In the Lab

The biggest part of the experiment will then be run in the lab. Students will be welcomed in groups of 4 to perform the task. We will first introduce them to the process, and then let them go through it on their own.

1. Pre-test

   First students will have to log in again to the website, so that they can fetch back the profile data they input previously. Then, they will have to perform a post-test task, cf. Figure 4. This task will provide Traditional material for both groups. We will record the time, but will not set a limit.

   We used less than 4 colours for the interface in order to not distract the user from the material. We have not put any colours in the materials, to mimic the real-life tests.

2. Practice

   Students will then be presented with 10 Practice tasks, that will have Traditional material for the Control group and Personalized material
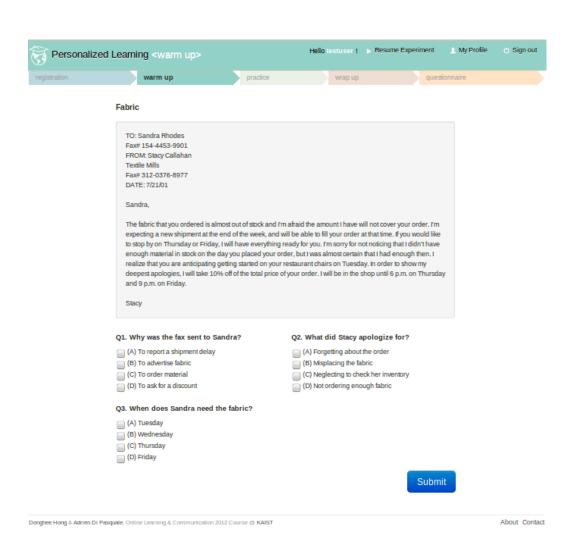
Figure 4: Pre test task

for the other. There will be no timing of this step, but they will have to complete in a certain amount of time. We will allow them to take breaks but no to have discussion among them, as to not bias the experiment.

3. Post-test

   This step is almost exactly the same as the Pre-test : Traditional material, time measured but not limited.

4. Questionnaire

   After the post test, we will ask students to take the questionnaire we wrote. Its purpose is mainly to evaluate how familiar they felt during practice, and measure their perceived engagement. It also includes two open feedback questions, that they can optionally fill.

5.2.3   Backend

The website is a complete platform, as it provides a complete backend to edit Tasks (cf. figure 5) and Questionnaires, and to handle Users (creation, step & material type assignment ..).

The forms can get a bit complicated because a big flexibility is required to input the Personalizable parts possible values, in the Tasks material, Questions texts and choices ... We will not get into details in this document but the website provides some on-screen indications, and more information can be found using the resources listed in the next section

5.3   External resources

The prototype source code and an online demo version can be found there :
`http://github.com/adipasquale/personalized_learning`
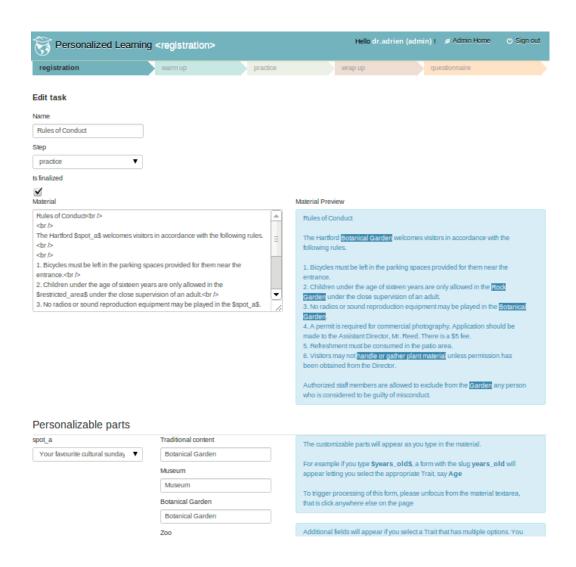It can freely be re-used and modified for any similar experiment.

Figure 5: Backend - Edit task

References

[1] C.-M. Chen, S.-H. Hsu, Y.-L. Li, and C.-J. Peng. *Personalized Intelligent M-learning System for Supporting Effective English Learning*, pages 4898–4903. Institute of Electrical and Electronics Engineers, Oct 2006.

[2] Y. Iijima and I. Horie. *A word list generator program for using authentic texts in an academic English reading class*, pages 407–412. Institute of Electrical and Electronics Engineers, Apr 2010.

[3] M. K. Khribi, M. Jemni, and O. Nasraoui. *Automatic Recommendations for E-Learning Personalization Based on Web Usage Mining Techniques and Information Retrieval*, pages 241–245. Institute of Electrical and Electronics Engineers, Jul 2008.

[4] C. L. López and H. J. Sullivan. Effects of personalized math instruction for hispanic students. *Contemporary Educational Psychology*, 16(1):95–100, Jan 1991.

[5] E. Mor and J. Minguillón. E-learning personalization based on itineraries and long-term navigational behavior. In *In Proceedings of the 13th international*, pages 264–265. ACM, 2004.

[6] J. C. Tseng, H.-C. Chu, G.-J. Hwang, and C.-C. Tsai. Development of an adaptive learning system with two sources of personalization information. *Computers & Education*, 51(2):776–786, Sep 2008.

[7] A. Türker, I. Görgün, and O. Conlan. The Challenge of Content Creation to Facilitate Personalized E-Learning Experiences. *International Journal on E-Learning*, 5(1):11–17, 2006.

[8] D. Xu and H. Wang. Intelligent agent supported personalization for virtual learning environments. *Decision Support Systems*, 42(2):825–843, Nov 2006.