



Chapitre d'actes

2015

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

## What Motivates Students to Use Online CALL Systems? A Case Study

---

Baur, Claudia; Rayner, Emmanuel; Tsourakis, Nikolaos

### How to cite

BAUR, Claudia, RAYNER, Emmanuel, TSOURAKIS, Nikolaos. What Motivates Students to Use Online CALL Systems? A Case Study. In: 9th International Technology, Education and Development Conference (INTED). [s.l.] : [s.n.], 2015.

This publication URL: <https://archive-ouverte.unige.ch/unige:46561>

# WHAT MOTIVATES STUDENTS TO USE ONLINE CALL SYSTEMS? A CASE STUDY

Claudia Baur<sup>1</sup>, Manny Rayner<sup>1</sup>, Nikos Tsourakis<sup>1</sup>

<sup>1</sup>University of Geneva (SWITZERLAND)

## Abstract

We present a case study using an online CALL game with secondary school students, where we examine the question of how different factors, such as gamification elements, difficulty of exercises and personal background can influence students' behaviour with regard to performance and motivation. For this study we used CALL-SLT, a speech-enabled online CALL tool for beginner learners of English. The content consisted of 8 interactive and dialogue-based lessons that were based on an English textbook commonly used in German-speaking Switzerland. The lessons are designed to help students practise their productive and receptive skills in a non-judgmental environment, where simple feedback (correct vs. incorrect) is provided in real-time by the computer. The tool was used in four secondary school classes (approx. 50 active users) all across German-speaking Switzerland for a duration of four weeks each. The experiment setup also required the students to take a placement test at the beginning and at the end of the experiment phase and to fill in a questionnaire on personal background at the beginning, as well as a qualitative feedback questionnaire at the end of the experiment. We examined three specific aspects with regard to their influence on motivation: (1) the influence of gamification elements, such as scores and performance-linked badges, (2) the influence of the exercises' difficulty and (3) aspects of students' personal background.

Keywords: CALL, Gamification, Motivational issues.

## 1 INTRODUCTION

Computer-Assisted Language Learning (CALL) is becoming increasingly popular in today's world. The advantages are manifold: CALL tools can help learners practice their skills in the language they are studying (the L2) at any time and any location, as long as they have access to a computer and, if relevant, a reliable broadband connection. Compared to other – classical – learning scenarios, the CALL approach to second language acquisition (SLA) promises a large saving in both time and money. Even though interaction with human native speakers cannot be replaced entirely, it can at least be imitated in a virtual environment, helping learners overcome their inhibitions against communicating in a foreign language. CALL makes it possible to offer exercises to language learners of all ages which let them practice both receptive and productive skills in a variety of situations.

So far most research projects in CALL have focussed on the added value that can be brought to language teaching and learning; in contrast, there is a lack of studies examining the motivation that causes students to use CALL tools in the first place. Here, we will be primarily concerned with the question of what makes students use (and enjoy using) online CALL tools. We describe a substantial experiment where we introduced CALL-SLT, a translation game with speech recognition, in two different schools in German-speaking Switzerland, and analyse factors affecting the motivation of students to engage in the game.

The paper is structured as follows. Section 2 gives some background on the CALL-SLT tool, section 3 describes the setup of the case study, and the results are presented in section 4. Section 5 concludes and gives pointers to future work.

## 2 CALL-SLT

CALL-SLT is an online CALL tool based on speech recognition, web and language processing technology [1]. It is a prompt-response system with multiple options for prompting the students. In the version presented in this paper, each prompt is a combination of a multimedia file in the L2 (here, English) and a written text instruction in the L1 (here, German). To give a typical example, the system plays a short cartoon clip with an English native speaker asking the question, "How many nights would you like to stay at our hotel?" and simultaneously displays the German text, "Frag: Zimmer für 6 Nächte" (*Ask: room for 6 nights*). The text indicates how the student is supposed to answer in the L2.

In this case, an acceptable response will be something like “I want a room for six nights”, “Do you have a room for six nights?” or “I would like to stay for six nights”. The intention is that a reasonably wide variety of grammatically and linguistically correct utterances will be accepted, as long as they correspond to the meaning of the German prompt.

Once the student has answered, by speaking into the headset, the system performs speech recognition and then matches the recognised utterance against the prompt’s specification of what should be counted as a correct answer; specifications are written in a simple formalism based on templates and regular expressions [2]. If there is a match, the system gives positive feedback by displaying a green frame around the text prompt, and moves on to the next dialogue state. If the utterance is rejected – this might be for a variety of reasons, including incorrect use of vocabulary, grammatical incorrectness, incorrect use of the user interface, bad pronunciation, bad recognition due to insufficient recording quality, etc. – a red frame (negative feedback) is shown and the student is asked to repeat or reformulate their response. Figure 1 shows the user interface.

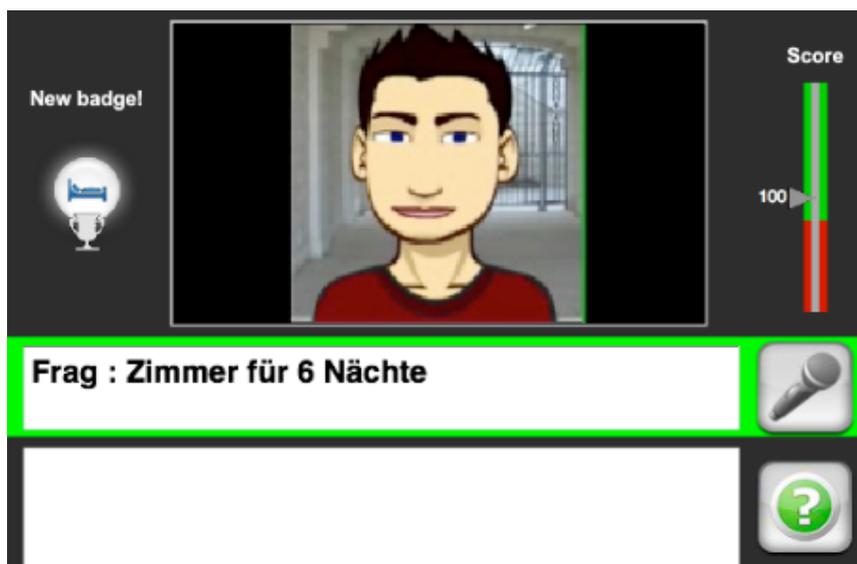


Figure 1: CALL-SLT user interface

## 2.1 Content

The content of the CALL-SLT course used for these experiments was developed in close collaboration with a Swiss secondary English teacher who supported us with his subject-matter expertise. We decided to create eight interactive lessons based on *Ready for English*, an English textbook commonly used in German-speaking Switzerland [3]. The eight lessons loosely cover the curriculum for the first year of English at secondary school. The lessons are the following: (1) *At the train station*, (2) *Getting to know someone*, (3) *At the tube station*, (4) *At the hotel*, (5) *At the restaurant*, (6) *Shopping* (7) *At the tourist information office* and (8) *Asking/Giving directions*; they can either be used independently or be put together to imitate a virtual trip to London. All lessons are dialogue-based with the goal of engaging the students in real life, situational conversations [4]. We focus on a communicative approach to SLA, putting more weight on achieving a successful interaction than on small grammatical or pronunciation flaws in the utterances.

## 2.2 Gamification Elements

The study used two different versions of CALL-SLT, one with gamification elements and one without. The gamified version of CALL-SLT (displayed in Figure 1) contains two main gamification elements that turn CALL-SLT into a serious game: a point-based score bar and a badge table. The students can gain or lose points during each interaction with the system: points are acquired if a complex response is given, and points are deducted if the utterance is not successfully recognised by the system. Every lesson starts off with an initial score of 100, depicted on a classical score bar. Users can potentially acquire four different badges for each lesson: plain, bronze, silver and gold. Plain and bronze badges only require the lesson to be completed a certain number of times (once for plain and three times for bronze). As well as completion, the silver and gold badges also require the student to achieve a minimum lesson score, with a threshold of 90 and 100 points respectively.

### 3 CASE STUDY

#### 3.1 Set-up

The experiment was carried out on four school classes at two different secondary schools in German-speaking Switzerland. The 49 active users were students between 14 and 16 years old. As mentioned in section 2.2, we used two slightly different versions of CALL-SLT that differed in two respects: feedback and gamification. The first version (standard) was the one described in [5]; the second version added a feedback module and gamification elements as additional features. The feedback module [6] gives the learners information both on technical problems (early release of the recording button, high background noise, improper recording quality) and also on their performance. The performance feedback is directly linked to the gamification elements, e.g. indicating how many more turns are needed to get the next badge. The main goal of comparing these two versions was to find out if the use of gamification elements in CALL tools helped motivate students (cf. section 4.1).

In order to get valid comparison results, we used parallel classes with the same teacher, meaning that we had two English teachers from two different secondary schools with two comparable classes each. The two paired classes each used a different version of CALL-SLT, one with gamification and one without. The versions were switched in the last week of usage in order to collect qualitative feedback from the students on both versions of CALL-SLT.

In class 1a we had 12 active users, in 1b there were 7 users, 2a had 21 users and class 2b had 9 active users. At the beginning of the experiment the users were asked to fill in a pre-experimental questionnaire, providing information on their personal background, and a placement test which evaluated knowledge of grammar and vocabulary. They were then asked to use the CALL-SLT tool autonomously as a homework assignment over a period of four weeks. At the end of the four weeks they filled in a post-experimental questionnaire giving feedback on their experience, and took another placement test that was identical to the first one.

### 4 RESULTS

We divided all classes into three motivation groups, for *high*, *medium* and *low* motivation respectively. The classification criteria were the number of logins and the number of logged interactions per student, giving information on both return rate (logins) and overall perseverance (interactions). In a second step we categorised each learner by recognition rate, depending on how often the student's answers were accepted by the system. Table 1 gives an overview of the applied criteria.

Low motivation		Medium motivation		High motivation	
Logins	Interactions	Logins	Interactions	Logins	Interactions
1	< 100	1-3	< 150	> 3	> 150
Low recognition rate		Intermediate recognition rate		High recognition rate	
< 50 %		50% - 79%		> 80%	

Table 1: Motivation annotation criteria

The motivation classification of all classes can be found in Table 2 and the respective recognition distribution is shown in Table 3.

	Low	Medium	High
Class 1a – no game	33.3%	50%	16.7%
Class 1b – game	57.1%	28.6%	14.3%
Class 2a – no game	-	-	100%
Class 2b – game	28.6%	23.8%	47.6%

Table 2: Motivation distribution

	Low motivation			Medium motivation			High motivation		
	Low	Int.	High	Low	Int.	High	Low	Int.	High
Class 1a	25%	-	8%	-	42%	8%	-	8%	8%
Class 1b	-	29%	29%	14%	14%	-	-	-	14%
Class 2a	-	-	-	-	-	-	-	67%	33%
Class 2b	5%	19%	5%	-	14%	10%	-	33%	14%

Table 3: Recognition distribution

The categories established in Tables 2 and 3 will be used to set the basis for all evaluations discussed in the remaining sub-sections. In 4.1 we examine how gamification elements influence motivation, 4.2 looks at the influence of the exercises' difficulty, and 4.3 gives an overview of how the students' personal background influences their motivation to use an online CALL game, such as CALL-SLT.

## 4.1 Gamification

In previous user studies, we noticed that motivation was a major issue in online CALL games. Our impression was that just being asked to use CALL-SLT as a homework assignment was not enough to get teenagers actually to use the tool, so we introduced the gamification elements described in section 2.2. In the present case study, we evaluate both the students' subjective perception of gamification elements, and objective measures like the number of acquired badges. As mentioned above, we evaluated two classes who used the gamified version and two classes using the standard version without gamification elements. In order to allow students of all four classes to give qualitative feedback on the two versions, we introduced the opposite version in the last week of the experiment phase, meaning that classes 1a and 2a used the standard version for the first three weeks and then the gamified version for one week, and classes 1b and 2b did the same in reversed order, starting with gamification and finishing the experiment with the standard version.

In the post-experimental questionnaire, all participating students were asked to give feedback on which version (gamified or standard) they liked more. As we can see in Table 4, all classes clearly judged the gamified version as their preferred alternative. When we split those figures up by the motivation categories established in Tables 2 and 3, we see that the proportion of low motivation students choosing the standard version was 25%. For medium motivation students, the figure was 63%, and for highly motivated students it was only 12%. This suggests that students who used the system more frequently appreciated gamification elements more than others.

	Gamified Version	Standard Version
Class 1a – no game	80%	20%
Class 1b – game	75%	25%
Class 2a – no game	86%	14%
Class 2b – game	90%	10%

Table 4: Gamified vs. Standard CALL-SLT

Although the subjective perception of gamification is very positive, we were surprised to find no objective connection between gamification and motivation. Going back to the motivation categorisation, we found that the proportion of highly motivated students in classes 1a/2a (standard version) was 52%, much higher than the 39% we observed in classes 1b/2b (gamified version). This in no way suggests that students with the gamified version are more motivated to use CALL-SLT.

An analysis of the number of badges collected by the four classes, as well as the distribution of the four badge levels, further undermines the claim that gamification motivates students from an objective point of view. Table 5 shows that classes using the standard version on average collected more badges than the two classes using the gamified version of the system, with a delta of 0.4 for 1a/1b and 3.0 for 2a/2b. (The software was configured so that students collected badges even when using the standard version, the difference being that in the gamified version the students were aware of collecting badges with the badges displayed in a badge table). This again fails to suggest quantitative evidence that the badges motivated students to use the system more.

The distribution of badges collected across the four difficulty levels (plain, bronze, silver and gold) does not present a clear picture either. Whereas in class 1b we have slightly more students who tried to get a more advanced badge than in 1a ( $\Delta$  9.5%), students of class 2b show the contrary behaviour

with more than 50% stopping at the plain badge level, compared to class 2a where more than 50% of all students (unknowingly) collected gold badges.

	Ø N° of Badges	Plain Badge	Bronze Badge	Silver Badge	Gold Badge
1a – no game	2.7	96.9%	3.1%	-	-
1b – game	2.3	87.50%	12.5%	-	-
2a – no game	7.6	8.2%	11.5%	29.5%	50.8%
2b – game	4.6	51.6%	17.5%	13.4%	17.5%

Table 5: Objective Badge distribution

In summary, the results above do not provide any obvious quantitative evidence that gamification increased motivation. However, since the subjective perception of gamification elements is highly positive across all classes, it would be interesting to conduct another study on the influence of gamification on motivation over a longer period of time and with larger exercise coverage, to see if the subjective appreciation of gamification influences the trend in the long run.

## 4.2 Prompt Difficulty

Another factor that we evaluated with regard to its influence on motivation is the difficulty of the prompts. We again used the information given in the post-experimental questionnaire as the subjective indicator. Table 6 presents the responses to the question of whether the students perceived the exercises' difficulty as *easy*, *medium* or *difficult*, again grouped into the three motivation categories established above. The results suggest that students generally considered the difficulty level to be adequate. However, students who considered the exercises to be too difficult tended to be less motivated to use the tool (25% of the less motivated students rated the exercises to be difficult, compared to only 14% of highly motivated students and 11% of the medium motivation group). Another conclusion that we can draw from these results is that students seem to be more motivated when they considered the difficulty of the exercises to be adequate for their current level of knowledge (68% in the high motivation group, compared to 50% in the low motivation group).

	Difficult	Adequate	Easy
Low motivation	25.0%	50.0%	25.0%
Medium motivation	11.1%	66.7%	22.2%
High motivation	13.6%	68.2%	18.2%

Table 6: Subjective prompt difficulty

The objective difficulty of the exercises was measured by determining the number of attempts needed per prompt. We again classified all turns in three difficulty categories: *easy*, *medium* and *difficult*. A turn was classified as *easy* if the student's response was accepted at the first attempt, and turns that were accepted by the system after the second or third attempt were classified as *medium*. The remaining prompts, which were not accepted after the third attempt and forwarded to a simple yes/no alternative, were classified as *difficult*. Table 7 shows the results grouped by motivation categories. We can see that the tendency is for easy and difficult prompts to be congruent with subjective perception (cf. Table 6). Across all classes, low motivation students had a smaller proportion of easy prompts than the students in the high motivation category, suggesting that the students were more motivated if the system accepted their responses after fewer attempts.

	Easy	Medium	Difficult
Low motivation	68.9%	26.6%	4.5%
Medium motivation	71.2%	24.5%	4.3%
High motivation	80.6%	16.0%	3.4%

Table 7: Objective prompt difficulty

Similarly, the perception of successful recognition also seems to have an effect on motivation. Even though the majority of all students subjectively considered the system to understand their answers well, the highly motivated students showed this behaviour in a more pronounced form than the other two groups (cf. Table 8). The objective recognition results further strengthen this conclusion, with a recognition rate of 60% in the low motivation group and 75% in the high motivation group. As with

prompt difficulty, students are more motivated when they get better recognition results from the system.

	<b>Good recognition</b>	<b>Bad recognition</b>	<b>Recognition Result</b>
Low motivation	61.5%	38.5%	59.5%
Medium motivation	66.7%	33.3%	69.7%
High motivation	72.7%	27.3%	75.2%

Table 8: Subjective and objective recognition results

### 4.3 Student's personal background

As a third potential indicator of motivation, we analysed the students' personal background: how much they liked studying English, their personal motivation, and their previous experiences with CALL tools and/or speech recognition technology.

Unsurprisingly, personal motivation seems to have a major effect on the students' return rate and persistence when using CALL-SLT, as can be seen in Table 9. Whereas half of the less motivated students regarded the use of CALL-SLT as an imposition because it was a homework assignment, only 29% of the medium motivation group and as little as 14% of the high motivation group shared this view. This suggests that if the students have a positive mindset towards CALL technology and think that doing the exercises is fun, as opposed to just having to do them as a homework assignment, they are in general much more motivated to use the tool.

	<b>Fun</b>	<b>Homework</b>
Low motivation	50.0%	50.0%
Medium motivation	71.4%	28.6%
High motivation	86.4%	13.6%

Table 9: Subjective Motivation

Table 10 gives the mean values of the subjective information the students gave in the pre- and post-questionnaires in answer to the question "How much do you like studying English?", where they could put a mark on a visual analogue scale (ranging from 1-10). We can see that all motivation groups indicated an increased score for their motivation to study English after the use of CALL-SLT. It is surprising that the medium motivation group had a lower overall score than the low motivation group. However, highly motivated students clearly seem to be the ones who (pre- and post-experimentally) are most motivated to study English with a pre-experimental score of 7.4, which was further increased to 7.8 post-experimentally.

	<b>Pre-experimental</b>	<b>Post-experimental</b>
Low motivation	7.1	7.4
Medium motivation	6.1	6.7
High motivation	7.4	7.8

Table 10: Scores "How much do you like studying English?"

As a last indicator for motivation, we analyzed the connection between previous experiences with CALL and/or speech recognition technologies. As the results in Table 11 show, students who previously used CALL tools were not more motivated to use CALL-SLT in our experiment. Prior exposure to speech recognition however seems to be relevant: 64% of all highly motivated students had already used speech recognition in the past, whereas only 36% of less motivated students had done so. This suggests that students who are knowledgeable about speech recognition and have realistic expectations of its power and limits are overall more motivated. Students who have never used speech recognition before might be frustrated more quickly if they are not correctly recognized and then be more inclined to abandon the game.

	<b>Previous knowledge of CALL</b>	<b>Previous knowledge of speech recognition</b>
Low motivation	42.9%	35.7%
Medium motivation	7.7%	53.8%
High motivation	22.7%	63.6%

Table 11: Previous experience with CALL and speech recognition

## 5 CONCLUSION

The evaluations discussed in this paper show some interesting (and partly surprising) tendencies. Despite the growing hype around gamification of CALL tools in the community, we did not see empirical evidence in our case study to strengthen the claim that students are more committed to CALL when gamification is involved. Nonetheless, our findings show that, across all classes and motivation groups, students subjectively liked gamification. Given the fact that the current experiment was of short duration (four weeks) and offered limited content (eight lessons), it would be interesting to extend the study, both in terms of time and content, and find out whether subjective preferences in favour of gamification influence objective measurements in the long run. The influence of prompt difficulty clearly confirmed our intuitive prediction that students are less motivated if the prompts are too difficult and if they do not get recognised often enough by the system. On a more subjective level, a positive attitude towards CALL and previous experience with speech recognition proved to help motivate students in online CALL exercises.

Another important motivation enhancer, which we did not analyse in this paper, is undoubtedly the teacher. It is the teacher's role to motivate the students to use the tool in the first place, which has an important effect on participation rate. The teacher also acts as a source of motivation during the experiment, by continuously encouraging the students to participate actively. On the other hand there is also the danger of provoking a halo effect if the students' sympathy or antipathy for the teacher falsifies the results. In future work on school-centred use of CALL, the teacher's role should certainly be examined in more detail.

## REFERENCES

- [1] Rayner, M., P. Bouillon, N. Tsourakis, J. Gerlach, M. Georgescu, Y. Nakao, and C. Baur. (2010). A multilingual CALL game based on speech translation. In Proceedings of LREC 2010, Valletta, Malta.
- [2] Rayner, M., C. Baur and N. Tsourakis. (2014). CALL-SLT Lite: A Minimal Framework for Building Interactive Speech-Enabled CALL Applications. In Proceedings of the WOCCI Workshop, Singapore.
- [3] Morrissey, F. A., Fäs, H., Marchini, D., and Stotz, D. (2001). Ready For English 1. Klett und Balmer AG, Zug, Switzerland.
- [4] Baur C., M. Rayner M. and N. Tsourakis. (2014). Crafting Interesting Dialogues in an Interactive Spoken CALL System. In Proceedings of EDULEARN, Barcelona, Spain.
- [5] Baur, C., M. Rayner and N. Tsourakis. (2013). A Textbook-Based Serious Game for Practicing Spoken Language. 6th International Conference of Education, Research and Innovation, (ICERI), Seville, Spain.
- [6] Tsourakis N., M. Rayner and C. Baur. (2014). Formative Feedback in an Interactive Spoken CALL System. In Proceedings of the WOCCI Workshop, Singapore.