

Is Computer Assisted Language Learning (CALL) efficient for grammar learning? An experimental study in French as a second language

Florence Chenu

Frédérique Gayraud

Bruno Martinie

Wu Tong

Laboratoire Dynamique du Langage, CNRS &
Université Lyon 2

Florence.Chenu@univ-lyon2.fr

The issue of the efficiency of CALL is controversial. Our goal is to investigate whether CALL has an advantage on the learning of French relative clauses by intermediate non-native learners. Participants were randomly distributed across a control and an experimental group. A pre-test confirmed that there was no significant difference between the two groups. The control group attended a 3-hour traditional class on French relative clauses. The same content was taught to the experimental group in a 3-hour self-training session during which participants used a CALL system. Finally, both groups took a post-test assessing their level of improvement. In addition, the experimental group had to fill out a questionnaire about their attitude towards CALL. Whereas most participants expressed a strong preference for traditional teaching, both methods revealed a similar improvement for both groups. Further analysis taking into consideration the level of proficiency revealed that low-level participants tended to make more progress in the experimental condition. We claim that classroom instruction involving both spoken and written presentations of the same material is detrimental to understanding, especially among less proficient learners.

Many studies have attempted to assess the impact of Computer Assisted Language Learning (CALL) on learning. These studies focus either on the achievement of linguistic skills such as speaking, listening, writing and reading (Garrett 1998) or on motivation (Skinner & Austin, 1999, Rico Garcia & Arias, 2000). Concerning grammar teaching, while some studies find a significant advantage of these techniques, others find no significant difference between CALL and other teaching methods. For instance, a positive impact of CALL has been found by Nutta (1998) investigating the effect of computer-based vs. teacher directed instruction on the acquisition of English as a second language. In this study, computer-based students reached better scores than teacher-directed students, leading the author to conclude that computer-based instruction can be an effective method of teaching L2 grammar. Similarly, Nagata (1996) investigated the efficiency of intelligent CALL (involving intelligent feedback) on the acquisition of Japanese postpositions by English-speaking students. One group of students was assigned to the Computer Assisted Language Instruction (CALI) condition and the other group took the workbook instruction condition. The experimental group obtained better scores in production than in comprehension. In addition, the students in the CALI group had significantly better attitudes towards their instruction than the students in the workbook group. The author concludes that the study demonstrates an advantage of an intelligent CALL program over a workbook instruction.

Other studies are less optimistic. Liou, Wang and Hung-Yeh (1992) compared two groups of Chinese speaking students learning English in two different learning conditions: one experimental group attended a ten-week CALL course while the control group was taught the same content in a paper-and-pen homework condition. On the basis of a questionnaire evaluating the subjects' attitudes, the authors conclude that the combination of classroom instruction and grammatical CALL might be helpful, "or at least not detrimental". But their survey does not demonstrate that CALL plus instruction would be more efficient than instructions plus homework. Chen (2005) observed the impact of traditional class instruction with or without computer aid on the acquisition of parts of speech (nouns, articles, pronouns, verbs, adjectives, adverbs, prepositions, coordination and subordination) by 2 groups of Taiwanese EFL learners. After a 16-hour instruction, both groups were asked to produce a written narrative. The dependent variable under consideration was the number of errors produced in each group and for each category of errors. Overall, there was no statistical difference between the control and the experimental groups. However, the experimental group performed better in the error categories of nouns and prepositions whereas the control group outperformed the experimental group for the error categories of lexicon and subject omission. The author fails to provide literature supporting those findings. Joy and Garcia (2000) underline that "much of the literature in the field of instructional technology purports to have found no significant difference in learning effectiveness between technology-based and conventional delivery media" (p. 33). According to the authors, much of this lack of significant difference can be accounted for by flaws in the experimental designs: for instance small number of subjects, absence of random assignment to control or treatment groups, small number of observations, no specification of the amount of time spent on the task.

Other factors are likely to affect the efficiency of the teaching method, and thus, possibly explain the above-mentioned contradictory findings. First, depending on the situa-

tion, grammar lessons can be taught in the native language of the learner, or in the target language. In the latter case, important comprehension skills in the foreign language are required and comprehension is precisely more demanding in a foreign language, due to the reduced size of available working memory in a foreign language (Berquist, 1997). Furthermore, foreign learners typically show a discrepancy between reading and listening skills (Diao & Sweller, 2007; Hirai, 1999; Wong, 2001). Reading comprehension shows an advantage over listening comprehension, especially in the early stages of foreign language learning. Thus another important factor in that situation is the level of proficiency of students: Hirai (1999) has shown that reading rates and listening rates are quite similar among highly proficient learners, but that listening comprehension is far behind reading comprehension among less proficient learners.

CALL and traditional situations may use different media. Even though Nagata (1996) and Dunkel (1991) for instance insist that the advantages of CALL that they observed are not due to the medium per se, other studies underline the impact of the channel of presentation, spoken, written, or both. Traditional instruction typically involves written and spoken channels. CALL systems can also use both or only rely on written language.

Several studies have shown a detrimental effect of simultaneous presentation of spoken and written material, especially in foreign language learners (Diao & Sweller, 2007). The present study aims at investigating whether CALL has an objective learning advantage on the acquisition of French relative clauses by intermediate non-native learners, taking into consideration the level of proficiency of the learners, and the impact of the medium (spoken + written vs. written only). We predict that less proficient learners will improve more using a CALL system resting on written presentation only.

Method

Participants

Twenty-six participants ranging in age from 18 to 42 years-old (14 native speakers of Mandarin Chinese, 4 of Japanese, 2 of Korean, 2 of Russian, 1 of Vietnamese, 1 of Spanish (Columbia), 1 of Hindi, and 1 of Azeri) were randomly distributed across a control and an experimental group. Both groups attended French classes at Lyon 2 University. The participants had been studying French for 19 months in average (no significant difference between the two groups) and had been living in France for 4 to 5 months in average (no significant difference between the two groups).

Procedure

In a pre-test, both groups were tested for their mastery of French relative clauses. The participants had to fill in the blanks with the appropriate relative pronouns (1) or to transform two independent clauses into a complex sentence containing a RC (2) (40 questions as a total).

- (1) La femme _____ il s'est marié est détestable¹.
- (2) Ça fait des semaines que je n'ai pas parlé avec ta mère. Ta mère ne m'apprécie pas².
Statistical analysis revealed no significant difference between the two groups ($t(24) = .346$, n.s.).

During the next phase, the control group attended a 3-hour traditional class, during which learners were taught the syntax of French relative clauses. The teaching included both grammatical explanations and practical exercises³. The lesson consisted in five parts: an introduction about RCs, subject RCs, direct object RCs, circumstantial RCs, indirect object RCs. The lessons consisted in examples and explanations. The explanations were given both in the written and spoken modalities. After each lesson, 2 or 3 exercises were proposed. The exercises were 7 to 14 sentences in length and were of different types: filling in the blanks, combining two independent clauses into a single sentence using a relative pronoun (or the other way round), multiple choice between relative pronouns.

In parallel, the same content was taught to the experimental group in a 3-hour self-training session during which participants used a CALL system specifically designed for the purpose of the study. This system was designed as a website including 5 lessons and 17 exercises showing an increasing difficulty. We insist that the content was exactly the same as in the control condition, but was delivered in the written modality only. Note that no intelligent feedback was provided by the system: after typing their answer, the participants were given the possibility to press a button to see the correct answer, without any further explanation. The participants in the experimental condition were only given the opportunity to access the site before the post-test. Finally, both groups took a post-test assessing their level of improvement. In addition, the experimental group had to fill out a questionnaire about their attitude towards their CALL experience.

Results

Efficiency

Table 1 presents the scores obtained by the experimental and control groups for the pre-test and the post-test.

For the pre-test, the difference between the two groups is not significant ($t(24) = 0.346$, n.s.). Similarly, the scores do not significantly differ between the two groups for the post-test ($t(24) = 0.497$, n.s.). For the control group, the difference between the pre- and post test is significant ($t(12) = 2.656$; $p=.02$), as well as for the experimental group ($t(12) = 2.970$; $p=.011$).

Next, we divided the participants according to the score they obtained in the pre-test: participants with a score under 15 out of 40 were qualified as 'less proficient' (13 partici-

1 The woman he married is dreadful

2 I have not spoken to your mother for weeks. Your mother does not like me

3 The material is available at <http://www.dcl-ish-lyon.cnrs.fr/annuaire/pdf/martinie/fle/sommaire.htm>

pants) while participants whose score was above 15 out of 40 were considered as 'highly proficient' (13 participants). In the pre-test, the less proficient group had a mean score of 9.80 out of 40 (SD = 3.186) while more proficient learners had a mean score of 18 out of 40 (SD = 3.490). In order to measure the improvement of each group of participants, we computed the difference between the pre-test and the post- test. Table 2 presents the results we obtained.

Table 1. Mean scores (and SD) in pre- and post-tests for the control and experimental groups.

	Pre-test	Post-test
Control group	14.61 (5.65)	17.73 (6.85)
Experimental group	13.84 (5.69)	19.07 (6.95)

Table 2. Mean difference between pre and post tests (and SD) in the control and experimental groups according to the level of proficiency.

	Low	High
Control group	1.2 (5.6)	4.3 (2.8)
Experimental group	7.3 (6.5)	1.9 (4.8)

This analysis revealed that low-level participants tended to make more progress in the experimental condition ($t(12) = 1.58, p=.14$): high-level participants gained 1.9 points whereas low-level participants gained 7.3 points. This result can not be accounted for by a ceiling effect in the more proficient group since their mean score was 18 out of 40 in the pre-test. By contrast in the control group, high level participants gained more points than low level participants, even if the difference does not reach statistical significance. In addition, we found an interaction of group by level of proficiency ($F(1,22) = 4.25 ; p=.05$).

Attitude

In addition, the experimental group was asked to fill out a questionnaire assessing their attitude towards grammar learning using CALL. The first series of questions aimed at evaluating the participants' involvement. 73% of the participants claimed that they read all the lessons, 66.7% made all the exercises 86.7% of which followed the order suggested by the site architecture. The participants were given the opportunity to access the site between the sessions, but only 3 out of them did so. Another set of questions aimed at assessing the perception of the level of difficulty of the lessons and exercises. 80% of the participants found that the lessons were "clear" and 20% found them "difficult to understand". Similarly, 73.3% of the participants found that exercises were "fairly easy" while 26.7% found them "difficult". 66.7% found the site user friendly, 26.7% found it "difficult to use" and 6.6%

gave no answer. 53.3% of the participants estimated that they were given enough time to read the lessons and to do the exercises. The last set of questions concerned the preference regarding teaching methods. 100% of the participants expressed their preference for traditional teaching, 86.7% of them explaining that the presence of a teacher allows more precise feed-back. In the same vein, 46.7% prefer to learn a language orally, 20% with written material and 33.3% think that both are suitable. Indeed many participants underlined that computer assisted language learning is appropriate for exercise practice in addition to a more traditional teaching.

Based on these questions, we computed an index of satisfaction and an index of difficulty. For the satisfaction index, points were given for each of the following questions: did you like the content of the lessons (yes = 1; no = 0)? Do you prefer to be taught grammar lesson by a teacher (0) or by a computer (1)? Do you prefer to learn a language in the spoken modality (0), in the written modality (2) both (1)? For the difficulty index, we asked the following questions: did you read all the lessons (yes = 0; no=1)? Did you do all the exercises (yes = 0; no=1)? Do you think the content of the lessons was very clear (1) clear (2) difficult (3) impossible to understand (4)? Do you think the exercises were easy (1) quite easy (2) difficult (3), very difficult (4)? Do you think you had enough time? (yes = 0; no = 1).

As can be seen from table 3, the level of proficiency of participants had no effect, neither on difficulty, nor on satisfaction.

Table 3. Mean satisfaction and difficulty indexes (and SD) in the experimental group according to the level of proficiency.

	Low	High
Satisfaction index	2 (1.95)	2 (1.22)
Difficulty index	5.8 (1.35)	5.6 (1.14)

Relations between attitudes and efficiency

Next, we analyzed the relations between the index of satisfaction and the index of difficulty and the scores obtained at the post-test. Globally speaking, no significant correlation was observed between subjective and objective measures ($p=.51$ for "satisfaction" and $p=.27$ for "difficulty").

Discussion

In our study the CALL condition has proved to be slightly more beneficial than traditional instruction for the learning of French RCs by non-native speakers. The experiment was designed in order to avoid the flaws of many experiments mentioned by Joy and Garcia (2000), even though more participants should be included in the study.

We obtained an advantage of CALL even though our system did not provide intelligent feedback. Nagata (1996) and Dunkel (1991) for instance insist that the advantages of CALL that they observed are not due to the medium *per se*, but that they are due to the quality

of the messages produced by the medium. However, in our study, the only difference between the two experimental conditions was the medium, which can be relevant especially if the level of proficiency is taken into consideration: Indeed, the benefit was more important for low-level participants. This result is consistent with Williams and Williams (2000) who observed an improvement with ESL low level students when they use a computer, even if the improvement was much greater in their study than in ours. This result sheds light on the importance of the modality *per se*, spoken or written. It may be a side effect of the status of writing in the two media: while traditional instruction implies both written and spoken language, the CALL device that we used only relied on written language. Written language lightens online processing constraints, enabling the learner to go at his own pace. Furthermore, using both written and spoken language produces a redundancy effect increasing the cognitive load to perform the task: when the same material is presented in different modalities, as it is often the case in traditional classroom instruction, learners must process two sources of information simultaneously, especially for less proficient learners for whom, decoding a text in one modality only taxes in itself working memory resources to a large extent (Diao & Sweller, 2007). Hence, CALL systems may not always be more efficient than traditional instruction, depending on the level of proficiency, and on the number of channels of presentation. In our case, we only used the written modality, but other CALL systems use dual mode presentations, because in multimedia educational environments, it has frequently been assumed that presenting the same material in both modalities would benefit learning and understanding (Kalyuga, Chandler, & Sweller, 2004). However, a growing body of research has shown a detrimental effect of such multimodal presentations (Kalyuga et al., 2004; Craig, Gholson, & Driscoll, 2002; Mayer, Heiser, & Lonn, 2001).

Hence, to compare efficiently traditional instruction and CALL, one should take care that both methods use the same modalities. If modalities of presentation are kept constant, the question then is to know whether a CALL instruction using only the written modality has an objective advantage over a traditional book. A possible difference might involve attitudes.

In our study, and contrary to what could be expected from the literature, no correlation between positive attitude and successful outcomes was observed. While a great emphasis is put on the relationship between positive attitudes and learning achievement in the literature, our study does not confirm such a correlation. In conclusion, our study shows that efficiency research should take into consideration factors such as the level of proficiency of learners and the modality of presentation.

References

- Berquist, B. (1997). Memory models applied to L2 comprehension: A search for common ground. In G. Taillefert & A.K. Pughs (Eds). *Reading in the university: First, second and foreign languages* (pp. 29-44). Toulouse: Presses de l'Université des Sciences Sociales.
- Chen, L. (2005). Examining the role of computers in EFL instruction. *Electronic Journal for the Integration of Technology in Education*, 4, 30-63.

- Craig, S.D., Gholson, B., & Driscoll, D.M. (2002). Animated pedagogical agent in multimedia educational environments: Effects of agent properties, picture features and redundancy. *Journal of Educational Psychology, 94*(2), 428-434.
- Diao, Y., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction, 17*, 78-88.
- Dunkel, P. (1991). The effectiveness research on computer-assisted instruction and computer assisted language learning. In P. Dunkel. (Ed.), *Computer-assisted language learning and testing: Research issues and practice* (pp. 5-36). New York: Newbury House.
- Garrett, N. (1988). Computers in foreign language education: Teaching, learning and language acquisition research. *ADFL Bulletin, 19*(3), 6-12.
- Hirai, A. (1999). The relationship between listening and reading rates of Japanese EFL learners. *Modern Language Journal, 83*, 367-384.
- Joy, E.H., & Garcia. F.E. (2000). **Measuring learning effectiveness: A new look at no-significant difference findings**. *JALN, 4*(1), 33-39.
- Kalyuga, S. Chandler, P., & Sweller, J. (2004). When redundant on-screen text in multimedia technical instruction can interfere with learning. *Human Factors, 46*, 567-581.
- Liou, H.C, Wang, S.H., & Hung-Yeh, Y. (1992). Can grammatical CALL help EFL writing instruction? *CALICO Journal, 10*(1), 23-44.
- Mayer, R., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*, 187-198.
- Nagata, N. (1996). Computer vs. workbook instruction in second language acquisition. *CALICO Journal, 14*(1), 53-75.
- Nutta, J. (1998). Is computer-based grammar instruction as effective as teacher-directed grammar instruction for teaching L2 structures? *CALICO Journal, 16*(1), 49-62.
- Rico Garcia, M., & Arias, F. (2000). A comparative study in motivation and learning through print-oriented and computer-oriented tests. *Computer Assisted Language Learning, 13*(4), 457-465.
- Skinner, B., & Austin, R. (1999). Computer conferencing: Does it motive EFL students? *ELT Journal, 53*(4), 270-277.
- Williams, H.S., & Williams, P. (2000). Integrating reading and computers: an approach to improve ESL students reading skills. *Reading Improvement, 37*(3), 98-101.
- Wong, W. (2001). Modality and attention to meaning and form in the input. *Studies in Second Language Acquisition, 23*, 345-368.

Biodata

Florence CHENU is Research Engineer at the French National Center for Scientific Research (CNRS) and member of the laboratory "Dynamique du langage" (UMR 5596 – CNRS). She received a PhD from Lyon2 University in 2003. She has taught CALL implementing at Lyon2 University for future teachers of French as a foreign language. Her research interests include first and second language acquisition. Her work has been published in *LIDIL*.

Frederique GAYRAUD is Assistant Professor of Psycholinguistics at the Université Lumière-Lyon 2 and member of the laboratory "Dynamique du langage" (UMR 5596 – CNRS). She received a PhD from Lyon2 University in 2000. She has taught French as a foreign language at Lyon2 University for over 10 years. Her research interests include first and second language acquisition. Her work has been published in *First Language, Journal of Psycholinguistic Research*.

Bruno MARTINIE is Assistant Professor of Linguistics at the Université Lumière-Lyon 2 and member of the laboratory "Dynamique du langage" (UMR 5596 – CNRS). He received a PhD from Paris X-Nanterre in 1998. He has taught French as a foreign language at Surrey University. His research interests include Syntax, Psycholinguistics and Second Language Acquisition. His work has been published in *Journal of Psycholinguistic Research* and *Recherches sur le Français Parlé*.

WU Tong is a Master's degree Student at Lyon2 University. His research interests concern the syntax of relative clauses in a typological perspective.