Design and empirical evaluation of controlled L2 practice through mini-games—moving beyond drill-and-kill?

Frederik Cornillie[1] and Piet Desmet[2]

Abstract. A key design issue for tutorial CALL is that controlled practice activities need to engage learners in meaningful L2 processing, so that any knowledge developed in such practice may transfer to meaningful L2 use in complex skills. Furthermore, activities for controlled practice ideally engender intrinsic motivation, so that learners are willing to practise and remediate problems outside of the classroom. This paper reports on an experimental and design-focused study that tackled these challenges by means of mini-games embedded in a mystery story read in class. Results show that intensive receptive practice helped learners to develop knowledge that was accurate and quickly retrievable on tests of near transfer, and that practice also transferred to more complex written productive tasks, as well as—to a smaller extent—to spoken productive tasks. We make suggestions for future design on the basis of the spoken productive language test used in this study.

Keywords: tutorial CALL, controlled practice, design, L2 grammar learning, mini-games.

1. ITEC, KU Leuven & iMinds, Kortrijk, Belgium; Frederik.Cornillie@kuleuven-kulak.be

2. ITEC, KU Leuven & iMinds, Kortrijk, Belgium; Piet.Desmet@kuleuven-kulak.be

1. Introduction

In current-day L2 pedagogy, the power of some kind of focus on form is undisputed, preferably in complex and meaning-oriented L2 tasks, disrupting the communicative flow to a minimal degree, and focusing on forms that are psycholinguistically relevant and necessary for the communication to succeed (Doughty & Williams, 1998). Yet, there is little scope in language teaching programmes for intensive controlled practice of specific linguistic constructions, accompanied by consistent Corrective Feedback (CF), equally embedded within meaning-oriented L2 use. Such practice may help to automatise knowledge in implicit memory, which could in turn free up attentional resources for higher-order skills during complex learning tasks. The state-of-the-art in second language acquisition theory assumes a dynamic interface between explicit and implicit knowledge (Ellis, 2005), and hence provides support for various attention-raising techniques, including controlled practice with CF. Furthermore, proponents of such practices have argued, from the perspective of skill acquisition theory, that CALL is *the* field that holds promise for the future of practice (DeKeyser, 2007), as it allows for massive and fine-grained data collection in longitudinal experimental designs, potentially in ecologically valid settings. Moreover, tutorial CALL affords the provision of consistent, error-specific, and supportive CF.

However, the implementation of tutorial CALL practice presents at least three serious design challenges. First, explicit focus-on-form practice needs to engage learners in meaningful L2 processing (DeKeyser, 1998; Wong & VanPatten, 2003), which is presumed necessary for realising transfer to complex skills. Secondly, given the little time there usually is for communicative L2 learning, teachers are likely to relegate practice to contexts outside of class. So, the key will be "to design interesting drills that are not demotivating" (Dörnyei, 2009, p. 289) and that—ideally—catalyse intrinsically motivated behaviour, so that learners are willing to practise without regulation from teachers. A third and related challenge is that consistent feedback inherent in practice may harm learners' competence needs (e.g. Robinson, 1991).

This paper reports on the empirical validation of an instructional design hypothesis for tutorial CALL practice that primarily addresses the first challenge (for the other two challenges, see Cornillie & Desmet, 2013), and which is inspired by principles from skill acquisition theory, task-based language teaching, and game design. To this end, we rely on the notion of *mini-games* (Cornillie & Desmet, 2016), i.e. gameful activities that can be completed in brief sessions, are
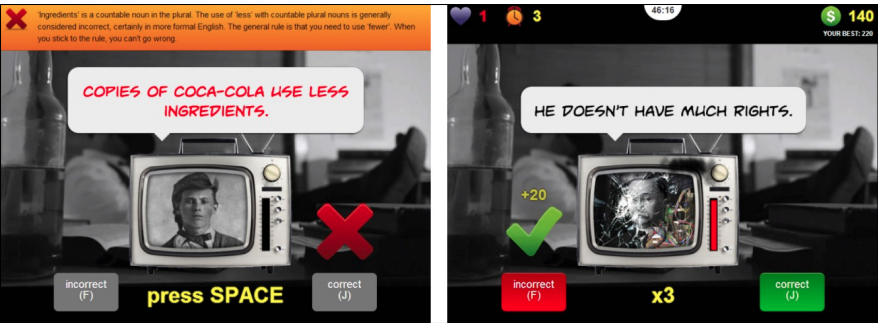
constrained in scope, provide consistent feedback, and thus lend themselves well to focused practice.

We address the following research question: to what extent does controlled L2 practice supported by mini-games and embedded in meaning-focused L2 use help learners to develop knowledge that is useful for their performance on various transfer tasks?

2.   Method

The study took place from January to March 2014 in secondary education in Flanders, and used an experimental between-subjects design with one control group (N=61) and two treatment groups (N=125). The treatment groups participated in reading and discussion activities concerning a mystery story based on the early history of Coca-Cola, in which mini-games were embedded which were intended to help the Dutch-speaking learners acquire the constraints on two grammatical features of English (quantifiers and the double object construction). Participants in the treatment groups received rule instruction, and were assigned at random to either a practice condition in which metalinguistic and error-specific CF was available (ML CF), or to a condition which only comprised 'knowledge of results' CF (KR CF), lacking metalinguistic explanation. Practice was receptive, and consisted of performing grammaticality judgments of sentences drawn from the mystery story; first in a 'tutorial' version of the mini-game, lacking time pressure and reward systems and comprising immediate CF for learners in the metalinguistic group (see Figure 1, left); then in a version involving time pressure, rewarding, between-learner competition, and vivid CF (i.e. 'knowledge of results' feedback adapted to the game fantasy) (see Figure 1, right).
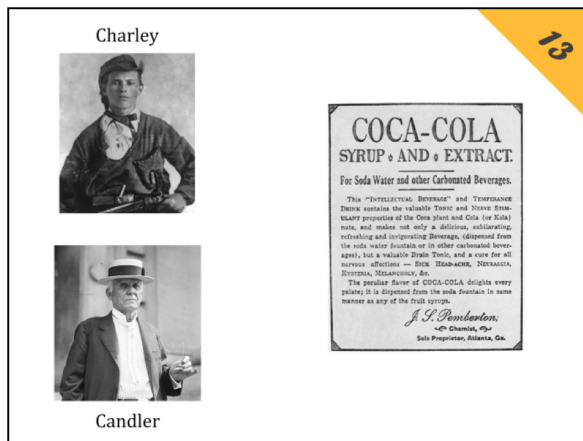
Figure 1. Tutorial version of the mini-game, with metalinguistic CF (left); full version of the mini-game (right)

The learners were introduced to the practice activities in class, and had opportunities to practise further at home. Practice behaviour was logged. The instructional procedure (i.e. instruction, text, and practice activities) lasted one month.

Prior to, immediately following, and one month after the procedure, all participants (including the control group) were tested on their knowledge of the target features using two transfer tasks: a Timed Grammaticality Judgment Test (TGJT; Loewen, 2009), and a Written Discourse Completion Test (WDCT). Participants in the treatment groups completed two more tests. First, they filled out a Metalinguistic Knowledge Test (MKT) aimed at measuring their knowledge of the grammar rules. Further, subsequent to the immediate post-tests, 69 learners were selected to participate in an oral production task known in the literature as the Oral Elicited Imitation Test (OEIT; Erlam, 2009). This test took the form of a role-play between the researcher and the participant, and required the learners to attend to both meaning and form. During the OEIT, participants were supported by means of slides to help them formulate their responses (see Figure 2). Learners were selected for this task on the basis of two parameters: the type of CF received during practice, and the amount of time spent on practice.
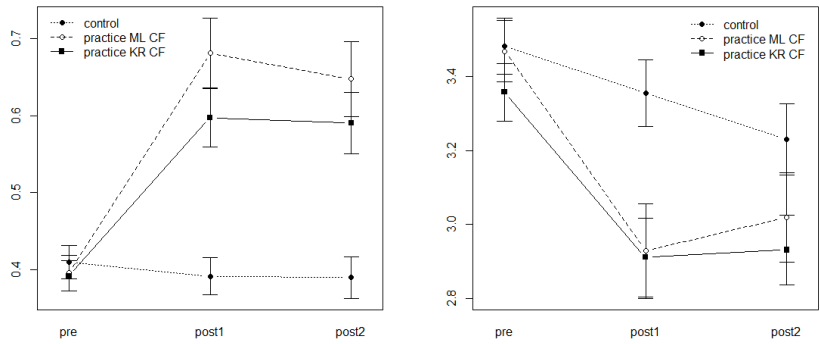
Figure 2. Visual support for the ungrammatical oral stimulus *Charley revealed Candler the secret recipe of Coca-Cola*, used in the OEIT
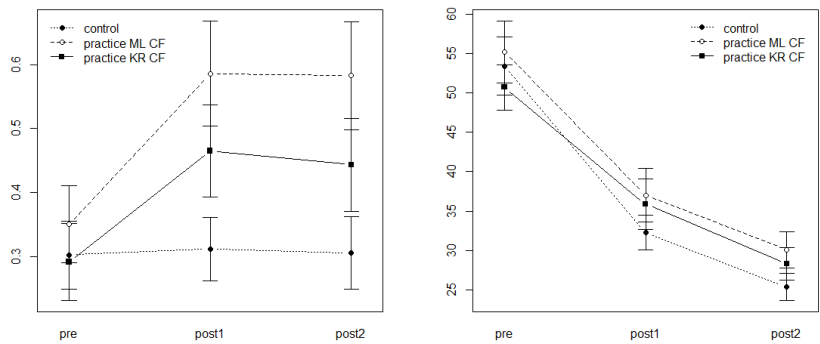

3.   Results

The data show that the treatment groups outperformed the control group on the post-tests of the TGJT in terms of accuracy rate and response times (see Figure 3).

Figure 3. Average accuracy rates (left) and response times (right) on the TGJT

As for the post-tests of the WDCT (Figure 4), the average accuracy rates of the participants in the treatment groups were higher than the average accuracy rate of the control group. Participants responded equally quickly in all groups.
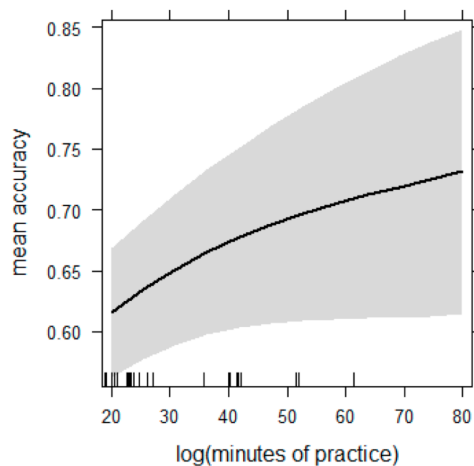
Figure 4. Average accuracy rates (left) and response times (right) on the WDCT

The linguistic accuracy scores on the OEIT were considered separately for the participants who had realised (despite the strong focus on meaning) that they were being tested on grammar (N=41) and for the unaware participants (N=8); the scores of the other 20 participants were disregarded, because it was unclear whether these learners had been focusing on form. The mean accuracy rates of the OEIT were regressed, for each group separately, onto two main predictors (i.e. feedback type; and time spent on practice, range between 2.9 and 85.7 minutes) and three control variables (accuracy scores on the pre-tests of TGJT and WDCT, and the MKT

scores). Two outliers were removed from the aware group, as these learners had misinterpreted the test instructions, affecting their scores negatively. The results of the regression analyses show that the mean accuracy rates of the aware group were positively affected by the time spent on practice (β=.084, p<.05) (see Figure 5) and by performance on the first WDCT (β=.322, p<.01). This regression model explained 32 percent of the variance in the mean accuracy rates (adjusted R^2=.32, $F(5, 32)$=4.523, p<.01). The same model applied to the unaware group revealed no effects.

Figure 5. Plot for relation between time spent on practice and mean accuracy on the OEIT (aware participants)

4. Discussion and conclusion

The results show that intensive practice with CF supported by mini-games and a mystery story helped learners to develop L2 grammar knowledge that was useful for their performance on various transfer tasks. There was evidence of transfer of practice to a follow-up task (TGJT) that was highly similar to the fairly simple and mechanical practice tasks (i.e. near transfer), but also to more complex written (WDCT) and spoken (OEIT) follow-up tasks (i.e. far transfer). Moreover, on the near transfer task, the knowledge developed during practice was quickly available.

However, observation of the learners in practice suggests that they were treating the practice tasks rather mechanically. Moreover, on the OEIT, learners were clearly

monitoring their spoken production, and the gains in accuracy were small. This may be due to the fact that the practice tasks were not very transfer-appropriate in relation to more complex tasks, and was consistent with skill acquisition theory, which posits that the effects of practice are skill-specific.

Therefore, if tutorial CALL practice aims to support learners in developing automaticity in speaking—which ideally it does—future research needs to find ways of engaging learners in more meaningful and productive, ideally spoken language practice. The OEIT, used in this study as a transfer test, is a primary candidate for such practice, especially if robust automatic speech recognition technologies can be used to elicit, structure, and give feedback on spoken language practice.

5.   Acknowledgements

References

Cornillie, F., & Desmet, P. (2016). Mini-games for language learning. In L. Murray & F. Farr (Eds.), *Routledge Handbook of Language Learning and Technology*. Routledge.

Cornillie, F., & Desmet, P. (2013). Seeking out fun failure: how positive failure feedback could enhance the instructional effectiveness of CALL mini-games. In *Global perspectives on Computer-Assisted Language Learning. Proceedings of WorldCALL 2013* (pp. 64-68). University of Ulster.

Cornillie, F., van den Branden, K., & Desmet, P. (2015). From language play to linguistic form and back again. Lessons from an experimental study for the design of task-based language practice supported by games. In J. Colpaert, A. Aerts, M. Oberhofer, & M. Gutierrez-Colon Plana (Eds.), *Task design and CALL; Proceedings of the Seventeenth International CALL Research Conference, Tarragona, Universitat Rovira i Virgili, 6-8 July 2015*. Antwerp: University of Antwerp.

DeKeyser, R. M. (1998). Beyond focus on form. Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). Cambridge: Cambridge University Press.

DeKeyser, R. M. (2007). Conclusion: the future of practice. In R. M. DeKeyser (Ed.), *Practice in a second language: perspectives from applied linguistics and cognitive psychology* (pp. 287-304). New York: Cambridge University Press.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Doughty, C., & Williams, J. (Eds.). (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

Ellis, N. C. (2005). At the interface: dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*(20), 305-352. doi:10.1017/s027226310505014x

Erlam, R. (2009). The elicited oral imitation test as a measure of implicit knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 65-93). Bristol: Multilingual Matters.

Loewen, S. (2009). Grammaticality Judgment Tests and the Measurement of Implicit and Explicit L2 Knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp, & H. Reinders (Eds.), *mplicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Bristol: Multilingual Matters.

Robinson, G. L. (1991). Effective feedback strategies in CALL: learning theory and empirical research. In P. A. Dunkel (Ed.), *Computer-assisted language learning and testing: research issues and practice* (pp. 155-167). New Jersey: Newbury House.

Wong, W., & VanPatten, B. (2003). The evidence is IN: drills are OUT. *Foreign Language Annals, 36*(3), 403-423. doi:10.1111/j.1944-9720.2003.tb02123.x

Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy
Edited by Francesca Helm, Linda Bradley, Marta Guarda, and Sylvie Thouësny