

Post-print version

This paper was published as:

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. Gutiérrez-Colón Plana (Eds.), *Proceedings of XVIIth International CALL Conference. Task design and CALL. 6-8 July 2015. Tarragona, Spain* (pp. 214–222). Antwerp: Universiteit Antwerpen.

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

Frederik Cornillie¹, Kris Van den Branden² & Piet Desmet¹

¹ KU Leuven – Faculty of Arts & iMinds - ITEC, Kortrijk, Belgium

² KU Leuven - Faculty of Arts & Centre for Language and Education, Leuven, Belgium

Introduction

In latest years, the field of CALL has witnessed a renewed interest in digital games. These are being seen to afford a heterogeneity of technology-mediated activities that catalyse ludic engagement in a second or foreign language (L2), very often in informal learning settings (e.g. Thorne, Black, & Sykes, 2009), which is typically highly meaning-focused, communicative, and oriented toward non-linguistic outcomes (Cornillie, Thorne, & Desmet, 2012). For designers of activities for instructed language learning, these characteristics of play in and with a L2 resound well with task-based language teaching (TBLT) (Ellis, 2003). However, in contrast with task-based language learning in informal contexts, such as online games, activities for instructed task-based learning call on a focus on form in order to help learners attend to aspects of the L2 that would otherwise pass unnoticed.

In TBLT, focus on form is typically realized in implicit ways, for instance by means of implicit recasts during interaction or by deferring feedback until the debriefing phase, in order to disrupt the communicative flow only to a minimal degree (Willis & Willis, 2007). Much less common is explicit attention to features of the L2 by way of form-focused exercises, i.e. pedagogical activities which do not result in a non-linguistic (communicative) outcome and are in this sense diametrically opposed to *tasks* in Ellis' (2003) model, but which are intended to help learners

develop understanding of a specific linguistic aspect. Furthermore, intensive controlled practice of specific linguistic constructions accompanied by consistent CF may help to automatize knowledge in implicit memory, which could in turn free up attentional resources for higher-order skills during complex learning tasks (Segalowitz & Hulstijn, 2005). Finally, extensive grammar explanation, for instance by means of metalinguistic CF, may help learners to realize transfer to different contexts of use and to different skills (DeKeyser, 1998).

The design of activities for controlled practice is challenging. First and foremost, given the available evidence that mechanical types of controlled practice do not promote and even hinder the development of communicative skills in a L2 (DeKeyser, 1998; Wong & VanPatten, 2003), activities for controlled practice need to engage learners in meaningful and ideally communicative L2 processing, similar to form-meaning processing in complex learning tasks. Secondly, controlled language practice demands consistent CF, tailored to individual errors and ideally extensive (e.g. by providing learners with metalinguistic explanation), which is hardly feasible in the language classroom. Third, 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.

To address these design challenges, we relied on the notion of *mini-games* (Cornillie & Desmet, forthcoming): gameful activities that can be completed in brief sessions, are constrained in scope, provide consistent feedback, and thus lend themselves well to focused – and potentially enjoyable – autonomous language practice. In an attempt to transcend mechanical drilling, the mini-games were embedded in a mystery story. The design hypothesis was that by interweaving the form-focused mini-games with meaning-focused reading and discussion activities, learners would be engaged in form-meaning processing during practice, promoting transfer of practice to more complex follow-up activities.

Our primary research question is: 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?

Method

The materials used in this study comprised a written text, mini-games, materials for grammar instruction, and four language tests. The text was a mystery story written by the first author on the basis of the early history of Coca-Cola, and served as the backdrop for reading and discussion activities in class. These activities created a context for meaningful practice, and were intended to help learners identify with the protagonist, a detective, and solve the mystery. In the parlance of TBLT, resolution of the mystery constitutes a *non-linguistic goal* for the learners.

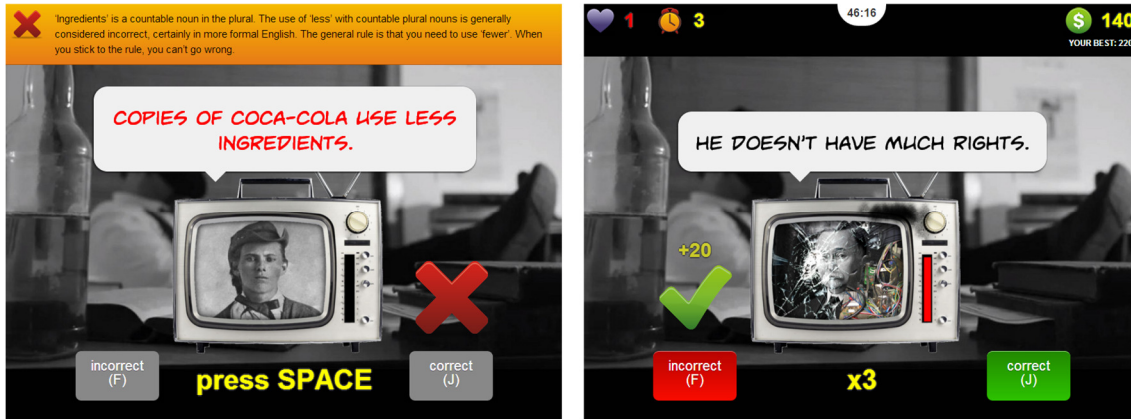
As a mechanism to advance the story, mini-games were inserted at two points in the text. These related to the content of the story on the one hand, and covered two grammatical problems of English on the other hand: quantifiers (QNT, e.g. *Charley has fewer shares in the company*) and the double object construction (V2O, e.g. **Pemberton revealed me the secret formula*). Importantly, (grammatical) instances of the double object construction in the mini-games also occurred in the reading text at later stages, which was meant to further automatize learners' knowledge of this problem while reading and discussing the text. Attempts were made to do the same for quantifiers, but these failed, as the text started to feel too artificial. To compensate for this, the first author tried, where possible, to use and elicit these constructions in the discussion activities.

Practice in the mini-games was receptive, and consisted of performing grammaticality judgments of sentences drawn from the mystery story. This activity was 'gamified' by adapting the format of the mini-game to the fantasy of the detective that interviews witnesses and potential suspects by means of a special interrogation device, by adding 'vivid' elements to

corrective feedback (the interrogation device broke down after 5 incorrect responses), and by including time pressure and a points system. In between the practice activities, the system displayed a *leaderboard* that compared the learners' individual best scores for the next exercise with the five highest (anonymized) scores.

The materials for grammar instruction comprised lists of sentences from which learners induced the grammar rules, followed by summary presentation of the rules provided by the first author and visually supported by slides. Grammar instruction for QNT took a conservative approach for the distinctions *fewer - less* and *fewest - least*, in line with learners' coursebooks. Instruction for V20 was largely based on Carroll and Swain's (1993) study.

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 ($N = 61$) to account for the potential effect of the tests and two treatment groups ($N = 125$). 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) in the mini-games, or to a condition which only comprised 'knowledge of results' CF (KR CF), lacking metalinguistic explanation. Next, the treatment groups participated in reading and discussion activities based on the mystery text, followed by practice with the mini-games. Learners first practised 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). In a second stage, learners engaged with a version that involved time pressure, rewarding, between-learner competition, and vivid CF (see Figure 1, right), followed by delayed grammar explanations for the metalinguistic group. The learners were introduced to the practice activities in class, and had opportunities to practise further at home. The system logged their practice behaviour. The instructional procedure (i.e. instruction, text, and practice activities) lasted one month.



**Figure 1: tutorial version of the mini-game, with metalinguistic CF (left);
full version of the mini-game (right)**

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.

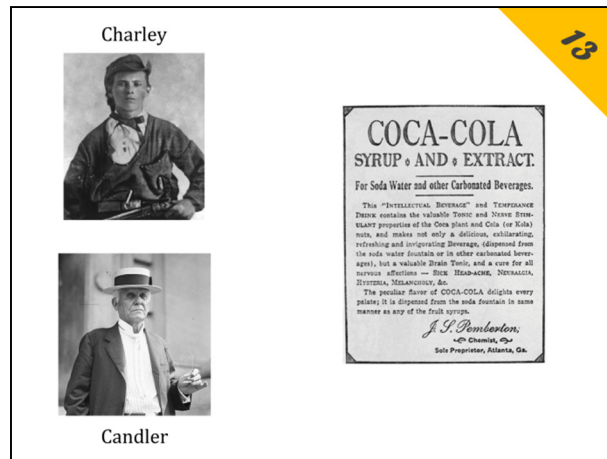


Figure 2: visual support for the oral stimulus
**Charley revealed Candler the secret recipe of Coca-Cola, used in the OEIT*

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, $p < .001$). Furthermore, the ML CF group outperforms the KR CF group in terms of accuracy on the first post-test ($p < .01$), but this effect disappears on the delayed post-test ($p = .069$).

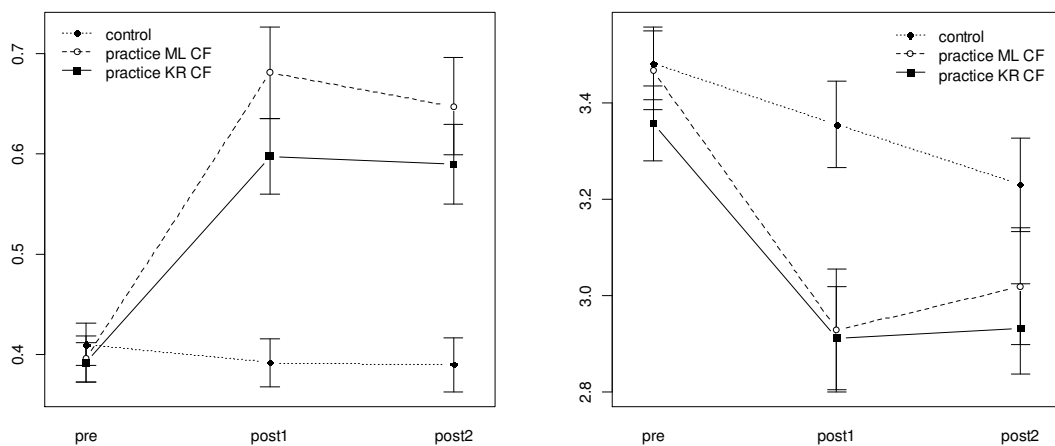


Figure 3: average accuracy rates (left) and response times (right) on the TGJT

As for the post-tests of the WDCT, the average accuracy rates of the participants in the treatment groups were higher than the average accuracy rate of the control group (all significant

at $p < .05$). The ML CF group outperforms the KR CF group at both post-tests (significant at $p < .05$). Participants responded equally quickly in all groups.

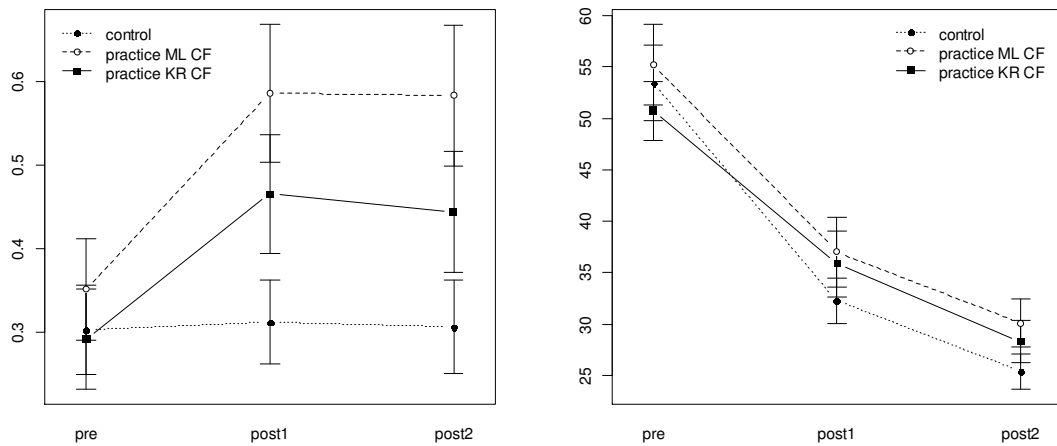


Figure 4: average accuracy rates (left) and response times (right) on the WDCT

Further worthy of mention are the interaction effects between time and linguistic problem (Figure 5). The effects of practice seem to be more durable for V2O.

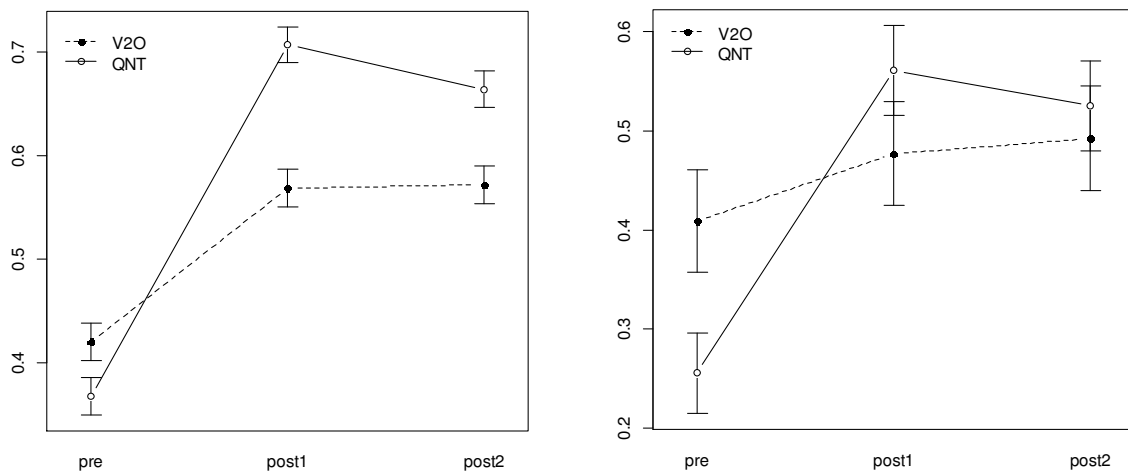


Figure 5: average accuracy rates for time x linguistic problem on TGJT (left) and WDCT (right) (treatment groups only)

As for the OEIT, the linguistic accuracy scores were considered separately for the participants who had realized (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 ($\beta = .084$, $p < .05$) (see Figure 6) and by performance on the first WDCT ($\beta = .322$, $p < .01$). 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.

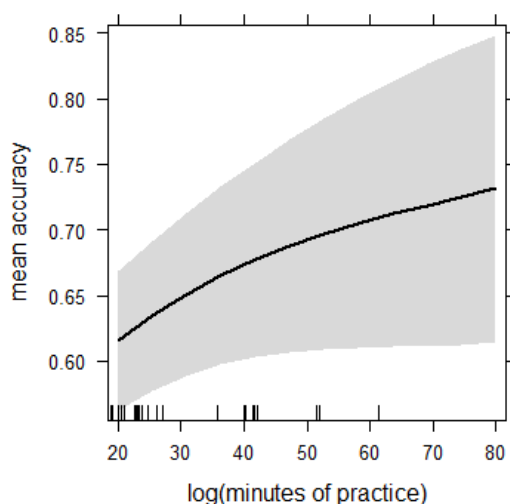


Figure 6: plot for relation between time spent on practice and mean accuracy on the OEIT (aware participants)

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 practice tasks (i.e. near transfer), but also to more complex written (WDCT) and spoken (OEIT) follow-up tasks (i.e. far transfer). The evidence of transfer from the

highly form-focused practice tasks to more complex follow-up tasks, which involved a stronger focus on meaning, may support the hypothesis that learners were engaged in meaningful language processing in the practice phase.

However, it is doubtful whether these effects can be attributed only to the controlled practice activities. First, observation of the learners in practice suggests that they were treating the practice tasks rather mechanically. Learners also had difficulties telling what the content of the items presented in practice were about. This indicates that they may not have been focused all that much on meaning during the activities for controlled practice.

Secondly, the effects of the treatment are puzzling when we consider the results for the linguistic problems. The effects for QNT decline on the delayed post-test; those of V20 do not, while this is clearly the more difficult grammar problem. One explanation may be that instances of the constructions for V20 were recycled more consistently during the reading and discussion activities – recall that it had proven impossible to systematically include instances of QNT in the mystery story.

Third, on the OEIT, learners were clearly monitoring their spoken production, indicating that they were switching to and fro between meaning focus and form focus. Moreover, the gains in accuracy are small. This may be due to the fact that the practice tasks were not very transfer-appropriate in relation to more complex tasks, and is consistent with skill acquisition theory, which posits that the effects of practice are skill-specific.

In conclusion, we suggest ideas for redesign of the materials. The challenge remains to engineer activities for controlled practice that are meaningful and communicative – in other words, activities that catalyse genuinely task-based language practice. A fairly simple trick on the level of the mini-games would be to add an activity that tests comprehension of the items that are presented in the grammaticality judgment tasks. Failure to respond to these comprehension questions could undo the points gathered while performing grammaticality judgments. Next, the potential of the mystery text for meaningful language processing could be

enhanced by truly involving learners as detectives in the story, rather than as readers who are supposed to identify with the main character. This could be done by creating interactive versions of the text with authoring tools for interactive fiction such as *Inform* (Aikin, 2009). Moreover, building in alternative ways of moving through the text could even create information gaps between learners, which can be exploited in pair work for productive communicative interaction. Finally, perhaps ways can be found to engineer mini-games for 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.

Acknowledgements

The conceptual design of the technology-based practice environment used in this study, but not its development, was partly realized through interaction with the Games Online for Basic Language learning (GOBL) project (519136-LLP-2011-NL-KA2-KA2MP), funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

The practice activity contains a still from the film noir project *The Big Smoke* (copyright 2012 Kenneth Gawne). Icons are from the Coquette icon set.

References

Aikin, J. (2009). *The Inform 7 Handbook*.

Carroll, S. E., & Swain, M. (1993). Explicit and implicit negative feedback. An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15(3), 357–386.

- Cornillie, F., & Desmet, P. (n.d.). Mini-games for language learning. In L. Murray & F. Farr (Eds.), *Routledge Handbook of Language Learning and Technology*.
- Cornillie, F., Thorne, S. L., & Desmet, P. (2012). Digital games for language learning: from hype to insight? *ReCALL*, 24(3), 243–256. doi:10.1017/S0958344012000134
- 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.
- Dörnyei, Z. (2009). *The Psychology of Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- 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 Teaching2* (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.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching* (pp. 94–112). Bristol: Multilingual Matters.
- Segalowitz, N., & Hulstijn, J. H. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. . B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). Oxford: Oxford University Press.
- Thorne, S. L., Black, R. W., & Sykes, J. M. (2009). Second Language Use, Socialization, and Learning in Internet Interest Communities and Online Gaming. *Modern Language Journal*, 93(0), 802–821.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Wong, W., & VanPatten, B. (2003). The Evidence is IN: Drills are OUT. *Foreign Language Annals*, 36(3), 403–423.