Learning L2 idioms in a CALL environment: the role of practice intensity, modality, and idiom properties

Catia Cucchiarini, Ferdy Hubers & Helmer Strik

Published online: 20 Apr 2020.

Submit your article to this journal ⬚

Article views: 3552

View related articles ⬚

View Crossmark data ⬚

Citing articles: 3 View citing articles ⬚

Learning L2 idioms in a CALL environment: the role of practice intensity, modality, and idiom properties

Catia Cucchiarini[a,b], Ferdy Hubers[a,b] (iD) and Helmer Strik[a,b]

[a]Centre for Language Studies, Radboud University, Nijmegen, the Netherlands; [b]Centre for Language and Speech Technology, Radboud University, Nijmegen, the Netherlands

ABSTRACT

Idiomatic expressions like *hit the road* or *turn the tables* are known to be problematic for L2 learners, but research indicates that learning L2 idiomatic language is important. Relatively few studies, most of them focusing on English idioms, have investigated how L2 idioms are actually acquired and how this process is affected by important idiom properties like transparency (the degree to which the figurative meaning of an idiom can be inferred from its literal analysis) and cross-language overlap (the degree to which L2 idioms correspond to L1 idioms). The present study employed a specially designed CALL system to investigate the effects of intensity of practice and the reading modality on learning Dutch L2 idioms, as well as the impact of idiom transparency and cross-language overlap. The results show that CALL practice with a focus on meaning and form is effective for learning L2 idioms and that the degree of practice needed depends on the properties of the idioms. L2 learners can achieve or even exceed native-like performance. Practicing reading idioms aloud does not lead to significantly higher performance than reading idioms silently. These findings have theoretical implications as they show that differences between native speakers and L2 learners are due to differences in exposure, rather than to different underlying acquisition mechanisms. For teaching practice, this study indicates that a properly designed CALL system is an effective and an ecologically sound environment for learning L2 idioms, a generally unattended area in L2 classes, and that teaching priorities should be based on degree of transparency and cross-language overlap of L2 idioms.

Introduction

Formulaic language refers to conventionalized word combinations that constitute an important part of vocabulary and are essential for native-like

CONTACT Catia Cucchiarini ✉ c.cucchiarini@let.ru.nl Centre for Language Studies, Radboud University, Nijmegen, the Netherlands

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

fluency and idiomaticity (Kecskes, 2007; Shei, 2008; Wray, 2002). Formulaic expressions are pervasive in native language (Pawley & Syder, 1983), but are less frequent in L2 and L2 learners experience difficulties in acquiring formulaic language (Ellis, Simpson-Vlach, & Maynard, 2008; Geluso, 2013; Nesselhauf & Tschichold, 2002). In addition, formulaic expressions, collocations and idioms are mentioned among the topics that still pose challenges to CALL research (Levy, Hubbard, Stockwell, & Colpaert, 2015). Collocations like *to commit suicide* or *heavy drinker* are known to be difficult for L2 learners and have been addressed in various CALL studies (Chan & Liou, 2005; Daskalovska, 2015; Futagi, Deane, Chodorow, & Tetreault, 2008; Nesselhauf & Tschichold, 2002; Rezaee, Marefat, & Saeedakhtar, 2015; Shei, 2008; Shei & Pain, 2000; Wu, Witten, & Franken, 2010). Idiomatic expressions like *to feel under the weather* or *to see eye to eye* also appear to be particularly challenging for L2 learners (Ahmadi, Sahragard, & Babaie Shalmani, 2017; Cieślicka, 2006; Conklin & Schmitt, 2008; Steinel, Hulstijn, & Steinel, 2007; Wray, 2000), but have received less attention in CALL.

Because research indicates that language is idiomatic in nature (Pawley & Syder, 1983), that learning idiomatic language positively influences L2 proficiency (Boers, Eyckmans, Kappel, Stengers, & Demecheleer, 2006; Hsu & Chiu, 2008) and that L2 learners consider learning L2 idioms important (Liontas, 2002), it is worthwhile investigating how L2 idiom learning can be facilitated. Previous research has shown how L2 idiom learning can be stimulated by employing web-based applications (Boers, Demecheleer, & Eyckmans, 2004; Boers, Eyckmans, & Stengers, 2007) or pedagogical agents (Ahmadi et al., 2017). Relatively few studies have investigated how idiomatic expressions are acquired in an L2 and how this process is affected by important properties of idioms like transparency (the degree to which the figurative meaning of an idiom can be inferred from its literal analysis) and cross-language overlap (the degree to which L2 idioms correspond to L1 idioms) as well as different types of practice. In addition, the majority of these studies have focused on English idioms (e.g., Ahmadi et al., 2017; Beck & Weber, 2016; Boers et al., 2007; Nordmann, Cleland, & Bull, 2014; Steinel et al., 2007; Titone & Libben, 2014). However, idiomatic expressions may incorporate specific linguistic and cultural knowledge (Boers et al., 2004; Kovecses & Szabco, 1996) and it is important to gather knowledge about idioms in languages other than English (Bonin, Méot, & Bugaiska, 2013; Caillies, 2009; Citron et al., 2016; Nordmann & Jambazova, 2017; Tabossi, Arduino, & Fanari, 2011).

The present study aims at filling in this knowledge gap by addressing the acquisition of Dutch idioms, which have not been studied to such a

degree before, and by adopting a twofold approach. First, we investigate the effects of intensity of practice and reading modality on learning idiomatic expressions in Dutch L2 by using a CALL system that provides German L2 learners of Dutch with immediate corrective feedback during practice. A CALL environment allows to systematically vary important variables such as practice intensity, modality of practice, and idiom properties. Second, we study the effect of transparency and cross-language overlap on learning Dutch idiomatic expressions.

Research background

Cross-language overlap

An important factor affecting L2 idiom processing is cross-language overlap, that is the degree to which L2 idioms have equivalents in L1. Irujo (1986) found that for advanced learners of English with L1 Spanish only identical idioms were easier to produce, while similar and different idioms were difficult because of L1 interference. In comprehension, on the other hand, the Spanish learners found L2 idioms that were identical or similar to L1 idioms easier, while idioms that were different were more problematic. Similar results were obtained for fluent Spanish-English bilinguals (Irujo, 1993).

A study on idiom avoidance by Laufer (2000) with Hebrew L1 learners of English L2 showed that idioms with partial similarity in form (E: miss the boat; H: miss the train) and idioms with no correspondence (E: it's not my cup of tea; H: no idiom) were avoided more than idioms with complete similarity in form (lay the cards on the table) and those with lack of similarity in form (E: to take someone for a ride; H: to work on someone) (Laufer, 2000, p. 186).

Charteris-Black (2002) made a clear distinction between conceptual similarity and linguistic similarity and found that idioms with equivalent concepts and equivalent linguistic forms were the easiest in comprehension (multiple choice task) and production (cued production task), while the most difficult ones were those with different concepts and equivalent linguistic forms and culture-specific idioms with different conceptual bases and different linguistic forms.

Türker (2019) recently showed that the presence of context might mitigate the effect of L1-L2 similarity. In a computer-mediated learning environment, English (L1) learners of Korean (L2) managed to learn L2 idioms irrespective of their degree of overlap with L1 idioms if sufficient context was provided.

Transparency

Transparency is generally defined as the degree to which the figurative meaning of an expression can be understood based on the meanings of its constituting words (e.g., Steinel et al., 2007) and is often measured by asking native speakers to indicate to what extent they "consider an idiomatic expression as related to its figurative meaning" (Skoufaki, 2008, p. 20). The idiom *kick the bucket* is opaque, because the figurative meaning to die cannot be extracted from the literal interpretation, while *to kill two birds with one stone* is transparent, as the figurative meaning (to complete two tasks by a single action) can be extracted from the literal interpretation. Transparent idioms appear to be less problematic for L2 learners than opaque ones both in idiom production and comprehension (Irujo, 1986; Skoufaki, 2008; Steinel et al., 2007; Yorio, 1989).

Steinel et al. (2007) investigated the effect of transparency on L2 idiom learning through a paired-associate learning (PAL) task in which Dutch university students were taught twenty English idiomatic expressions varying in transparency. High transparent idioms appeared to be comprehended better than low or intermediate ones. In order to gain more insight into the notion of transparency Skoufaki (2008) presented advanced L2 learners of English with unknown idiomatic expressions, varying along the transparency dimension, and asked them to guess the meaning and provide an interpretation. She found that high-transparency idioms received fewer different interpretations than low-transparency idioms, which she interpreted as evidence for the fact that it was easier for L2 learners to comprehend transparent unfamiliar idiomatic expressions than non-transparent unfamiliar ones.

Types of practice in idiom learning

Boers and Lindstromberg (2012) present an insightful categorization of instructional interventions and their effect on the acquisition of L2 formulaic expressions. Most of the available studies addressed the acquisition of collocations, though some investigations on idiom learning are also reported. What is remarkable about this review and about L2 idiom research in general is that relatively few studies have investigated how actual practice with idiomatic expressions affects their learning.

For example, little is known about the effects of practice intensity on L2 idiom learning. A study on collocations by Durrant and Schmitt (2010) showed that after two exposures recall of 20 adjective-noun collocations was successful in five out of 10 posttest items, while Webb, Newton, and Chang (2013) found that more encounters led to better recall, but that even after 15 encounters no full scores were obtained.

Boers and colleagues used a web-based application to provide different types of practice on idiomatic expressions. Boers et al. (2004, 2007) showed that awareness of the etymology of English idioms helped L2 learners remember those expressions. Boers, Lindstromberg, Littlemore, Stengers, and Eyckmans (2008) showed that pictorial elucidation contributed to better retention of idiom meanings, but did not help recollect their exact linguistic forms. The authors hypothesized that images might actually distract learners from paying attention to the linguistic forms, which was confirmed in Boers, Píriz, Stengers, and Eyckmans (2009). In a more recent study, Stengers, Deconinck, Boers, and Eyckmans (2016) modified their web-based learning application in an attempt to engage L2 learners with the linguistic form of L2 idioms through idiom copying. In addition to the online exercises adopted in previous studies, half of the Dutch (L1) learners of English (L2) participating in this study were asked to type each idiomatic expression in a type-in window (an exercise drawing attention to form), while the other half rated the usefulness of each idiom on a 5-point scale (a task that did not focus on form). The copying exercise proved not to be effective in stimulating retention of the linguistic form of the expressions. The authors provide several explanations for these results, including the limited congruence of the copying exercise (which did not require retrieval from memory) with the test (which did appeal to retrieval ability), and the relative shallow form of engagement induced by copy-pasting.

This research overview indicates that several important insights have been obtained on L2 idiom processing and learning, but that many questions remain. For instance, research has shown that practice with idiomatic expressions is beneficial for learning, but it is not clear how intensive this practice needs to be and whether intensity of practice should vary for idioms that vary along different dimensions, such as transparency and cross-language overlap. Furthermore, previous studies show that practice should focus on both meaning and form (Boers et al., 2008, 2009), but do not specify how this could best be done. The review also shows that cross-language overlap is an important property affecting L2 idioms processing, with idioms that are identical in L1 and L2 being easier to learn than idioms that are different. However, it is less clear how intermediate levels of cross-language similarity affect L2 idiom learning. With respect to transparency, research suggests that transparent idioms pose fewer problems to L2 learners (Irujo, 1986; Skoufaki, 2008; Steinel et al., 2007; Yorio, 1989), but there is little information about the amount practice required to learn idioms of different transparency levels. Finally, for these important idiom properties like cross-language overlap, transparency and frequency of exposure, little is known about

which L2 idioms should be prioritized in teaching and how this should be established.

The present study

We conducted a CALL-based study examining the effects of practice intensity and reading modality on L2 idiom learning, as well as the impact of important idioms properties like transparency and cross-language overlap.

An additional innovative element of this study is that it is couched in a broader research program investigating learning, representation, and processing of formulaic language in L1 and L2: the Idiomatic Second Language Acquisition (ISLA) programme[1], which has important benefits. First, we can rely on a large database of normative native data on idiom properties and idiom knowledge (see Hubers, van Ginkel, Cucchiarini, Strik, & Dijkstra, 2018, Hubers, Cucchiarini, Strik, & Dijkstra, 2019) especially compiled for this program by collecting data on 374 idioms from more than 394 native Dutch (L1) participants. This database is used in the present study to select idiomatic expressions for pre-tests, post-tests and learning experiments, which allows us to select the idioms based on normative data for important idiom properties, such as cross-language overlap, frequency of exposure, familiarity, transparency, image-ability and the extent to which each expression is actually known by native speakers.

Second, a CALL system is part of the ISLA program to which this study is linked. The use of a CALL system in our learning experiment ensures the ecological validity of experimental conditions, in the sense that they could be easily reproduced in teaching practice if the treatment turns out to be successful. Using the CALL system, we can create conditions that favor explicit learning of idiomatic expressions by L2 learners through more intensive practice than would be possible in traditional teacher-fronted contexts.

This study builds on previous research that employed CALL-based approaches (Boers et al., 2004, 2007; Türker, 2019) and reinforcement activities with a focus on form (Stengers et al., 2016). As explained by Stengers et al. (2016), the copying activity they applied was not sufficiently effective in drawing attention to form. In this study, we investigate two alternative ways of drawing attention to the linguistic form of the idiomatic expressions: two types of a reading activity, reading aloud and silently. The rationale behind this choice is that reading aloud should involve more engagement with the phonological form of the

idiomatic expression, which is likely to leave stronger memory traces and therefore could lead to better retention.

Research questions and hypotheses

In our study, we addressed the following research questions:

1. To what extent does practice intensity contribute to L2 idiom learning?
   We predict that practice intensity will have a positive effect on learning, with better performance on idioms that are practiced more intensively.
2. To what extent is the effect of practice related to idiom properties such as transparency and cross-language overlap?
   We expect the effect of practice to be modulated by idiom properties such as transparency and cross-language overlap. In particular, based on findings related to the effect of transparency on L2 idiom learning (Skoufaki, 2008; Steinel et al., 2007) we hypothesize that transparent idioms require less intensive practice than opaque idioms . Concerning the effect of cross-language overlap, we expect idioms that do not exist in L1 to be more difficult to learn (Irujo, 1986). However, based on available research it is difficult to predict how intermediate degrees of cross-language similarity will affect learning. A partial overlap may be expected to have a facilitative role, but it could just as well be confusing to L2 learners.
3. Does reading modality (reading silently or aloud) during practice have a differential effect on L2 idiom learning? We hypothesize that reading aloud should be more conducive to learning than reading silently for the simple reason that actively pronouncing the words in the idioms requires more phonological processing and possibly attentional capacity than reading them silently.

Method

Participants

The participants in this study were 42 German L1 learners of Dutch L2 who studied at a University of Applied Sciences in Nijmegen (HAN). Their mean age was 21 (SD = 2.1) and their proficiency level was intermediate, B1 according to CEFR. To obtain more objective data on their proficiency level in Dutch, we administered the Dutch version of the LexTale test (Lemhöfer & Broersma, 2012). The average score turned out

Table 1. Mean scores and SDs by native speakers for the idiom properties (scale 1-5) and idiom knowledge (% correct) of the idiomatic expressions included in our experiment and for the subsets A and B.

	Complete dataset (n = 60)	Set A (n = 30)	Set B (n = 30)
Frequency	3.43 (0.78)	3.42 (0.80)	3.44 (0.77)
Familiarity	3.17 (0.88)	3.18 (0.65)	3.15 (0.86)
Usage	2.08 (0.72)	2.08 (0.91)	2.07 (0.79)
Transparency	2.87 (0.75)	2.82 (0.74)	2.93 (0.77)
Imageability	3.41 (0.80)	3.23 (0.82)	3.60 (0.75)
Idiom knowledge	79 (0.23)	79 (0.23)	79 (0.23)

to be 64.7 (SD = 8.3). This study was ethically assessed and approved of by the Ethics Assessment Committee (EAC) of our faculty.

Material

Idiomatic expressions
We selected 60 Dutch idiomatic expressions from our native benchmark database consisting of 374 idioms with scores about idiom knowledge and various idiom properties (Hubers et al., 2019, 2018). We selected expressions that varied in terms of transparency, and cross-language overlap with German, and that were generally known by native speakers, as we thought this was an important criterion for teaching L2 idioms: It would be odd to teach L2 learners idiomatic expressions that are not even known by native speakers. See Table 1 for the mean scores of the idiomatic expressions selected on important idiom properties. The idiomatic expressions included in the experiment are presented in the Appendix, Table A1.

For cross-language overlap, we decided to adopt a four-level classification similar to those applied by Laufer (2000) and Titone, Columbus, Whitford, Mercier, and Libben (2015): (1) The Dutch idiom does not exist in German (NE), (2) The Dutch idiom does exist in German, but in completely different content words (DW), (3) The Dutch idiom does exist in German and has n content words in common (nW), (4) The Dutch idiom has a word-to-word correspondent in German (AW). Two German-Dutch bilingual students determined cross-language overlap separately and subsequently compared their scores. Potential differences were resolved by discussing these idioms in more detail.

Learning exercises
After consultations with teachers of Dutch L2, we decided to design exercises according to an approach that is widely used in DL2 teaching, the ABCD model by Neuner, Krüger, and Grewer (1981). Four steps are envisaged in this approach:

A. Input of new language elements through reading and/or listening
B. Guided processing through gap-fill or matching exercises
C. Semi-guided processing through answering questions or finishing sentences
D. Conversation or discussion, role playing or other more open exercises with a teacher

Steps A-C can be facilitated in a CALL environment, while D can best be practiced with a teacher (Elshoff, 2014). In line with these principles, we developed the following exercises: (1) PAL paradigm (step A), (2) a gap-fill exercise (step B), (3) sentence completion (step C), and (4) idiom selection (step C).

In the PAL paradigm the idiomatic expressions were presented to the participants along with their meanings in a one-by-one fashion. As part of the gap-fill exercise, participants were presented with sentences containing an idiomatic expression in which one word was missing and were asked to provide the missing word by typing. These sentences were constructed in such a way that they biased the reader towards an idiomatic interpretation. For the sentence completion exercise, participants were prompted with comparable idiomatically biasing sentences and were presented with three possible continuations containing idiomatic expressions. They were asked to type the appropriate continuation. In the idiom selection exercise, participants were presented with the meaning of an idiom and were asked to select the idiomatic expression out of three possible idioms that corresponded to the given meaning. See the Appendix for screenshots of the various exercises (Figures A1–A4). For the gap-fill, sentence completion, and idiom selection exercises, participants received corrective feedback after typing in the answer, indicating both the correct answer and the answer as provided by them (see screenshot in the Appendix, Figure A5). Subsequently, participants were presented with the context sentence including the correct answer (in case of gap-fill and sentence completion exercises) or with the correct idiom in isolation (in case of the idiom selection exercise). Depending on the condition the idiom was assigned to participants had to read aloud the prompt (as indicated by a record button) or had to read it silently (if record button was absent). A screenshot of both situations is included in the Appendix (Figure A6 and A7 respectively). For each idiomatic expression three context sentences were created, in order to prevent participants from associating an idiom to a specific sentence.

Design and procedure

We adopted a pre-test post-test within subjects design with the variables Intensity of Practice (Intensive = 12 times and Limited = 2 times) and

Reading Modality (Silent and Aloud). The effect of Reading Modality was only assessed for idioms that received intensive practice.

In order to assess the effect of Intensity of Practice, the idiomatic expressions were divided into two sets, A and B, of 30 expressions. These sets of idioms were alternately used in the intensive and limited practice conditions. These sets were balanced with respect to idiom properties as assessed through subjective judgments of frequency (of exposure), usage, familiarity, imageability and transparency based on the following questions (Hubers et al., 2019, 2018):

1. Frequency: How often have you heard or read this expression? (1. very rarely – 5. very often)
2. Usage: How often have you used this expression yourself? (1. very rarely – 5. very often)
3. Familiarity: How familiar are you with the meaning of this expression? (1. completely unfamiliar – 5. completely familiar)
4. Imageability: How easily can you form an image of this expression? (1. very hard – 5. very easily)
5. Transparency: How clear is the meaning of this expression based on the individual words in the expression? (1. very unclear – 5. very clear)

See Table 1 for an overview of the mean scores and SDs on these dimensions (scales 1 – 5) for all idiomatic expressions included in this experiment, and for the idioms in subsets A and B.

As mentioned above, Reading Modality was only assessed in the intensive practice condition, that could be either the idioms in set A or the idioms in set B. In order to test the effect of Reading Modality, sets A and B, were again divided into two, resulting in sets A1 and A2, and B1 and B2. These subsets contained 15 idiomatic expressions and were all selected in such a way that it was a plausible reflection of complete dataset with respect to idiom properties.

The materials were divided in four master lists in such a way that all materials were counterbalanced. See Table 2. We created three versions of each master list, counterbalancing the context sentences over the training sessions. In that way, the same context sentence was not presented in the same exercise in the same training session for all participants. Each participant received one of the 12 lists.

The experiment was divided into five sessions: a pre-test, three training sessions using the CALL system, and a post-test. All sessions were located in a computer room where the participants had access to a computer with internet and head set.

Table 2. Distribution of the materials over de master lists.

	Intensive practice		
	Silent	Aloud	Limited practice
Master list 1	B1	B2	A1, A2
Master list 2	B2	B1	A1, A2
Master list 3	A1	A2	B1, B2
Master list 4	A2	A1	B1, B2

Pre-test

The pre-test was conducted online through the Qualtrics platform (Qualtrics, Provo, UT) and included a background questionnaire. As part of a larger study on Dutch idiomatic expressions participants also rated each Dutch idiom by answering the five questions presented above on frequency, familiarity, usage, imageability and transparency. In addition, their idiom knowledge was tested in a multiple-choice question, in which participants had to select the correct idiom meaning out of four alternatives, and in an open-ended question, in which participants were asked to type in the corresponding meaning. For more detailed information about this procedure, see Hubers et al. (2019, 2018). Participants ended the pre-test by filling in the Dutch version of the LexTale, a test to measure vocabulary knowledge (Lemhöfer & Broersma, 2012).

Training

The training sessions were conducted online through a CALL system. During the first training session, participants were presented with the PAL paradigm that included both the idioms that received intensive and limited practice. Subsequently, they completed the gap-fill exercise only for the intensive practice idioms. The first training session took about 45 minutes. The second and third training session both consisted of the sentence completion exercise and the idiom selection exercise for the intensive practice idioms only, and took on average one hour.

Post-test

The post-test consisted of two parts: a sentence completion exercise without feedback, and a reduced version of the pre-test. Participants started with a sentence completion exercise that included all idiomatic expressions and that was conducted in the CALL environment. Subsequently, their idiom knowledge of all idioms was tested by means of the multiple-choice and open-ended questions also included in the pre-test through the Qualtrics platform (Qualtrics, Provo, UT). It took the participants on average 75 minutes to complete the post-test.

Table 3. Mean proportions correct and SDs on the multiple choice question.

|	Intensity of Practice	
Test	Limited	Intensive
Pre-test	0.55 (0.50)	0.56 (0.50)
Post-test	0.65 (0.48)	0.83 (0.38)

Mean native performance: 0.79.

Data analysis

To examine to what extent Intensity of Practice, Transparency, Cross-Language Overlap and Reading Modality affected idiom learning, we conducted two logistic mixed effects regression analyses. The statistical software package 'R', version 3.4.0 (R Development Core Team, 2008), was used to analyze the data, and the R packages 'lme4' (Bates, Mächler, Bolker, & Walker, 2015), 'lmerTest' (Kuznetsova, Brockhoff, & Christensen, 2017), and 'effects' (Fox, 2003). The regression models were built in a forward manner, starting off with a base model including our variables of interest (Test, Frequency of Practice, Transparency, and Cross-Language Overlap) and a random effect of Participant (random intercept only). The variable Reading Modality was analyzed in a separate regression analysis because this variable is only relevant for a subset of the data (idioms that received intensive practice). The basic model for this analysis included Reading Modality, and Test as fixed effects and Participant as a random effect (random intercept only). Subsequently, we added fixed and random effects one-by-one based on theory and examined whether the model fit improved. If the model fit did not improve, the predictor was not included in the model. The final models are reported in this paper.

Results

Intensity of practice, transparency, and Cross-Language overlap

Table 3 shows the mean proportions correct on the multiple choice question targeting idiom meaning recognition split by Test and Intensity of Practice. In the post-test participants performed better than the native speakers in Hubers et al. (2019) on the idiomatic expressions that had received intensive practice. See Figure 1 for a visualization of the data.

In order to examine the role of Intensity of Practice, and Transparency and Cross-Language Overlap in idiom learning, we carried out a logistic mixed effects regression analysis. The effect of Reading Modality was assessed in a separate analysis, since this was only relevant for a subset of the data, i.e., the idioms that received intensive practice. The responses to the multiple-choice question were converted to a binary variable (correct/incorrect) and used as the dependent variable in the analysis.
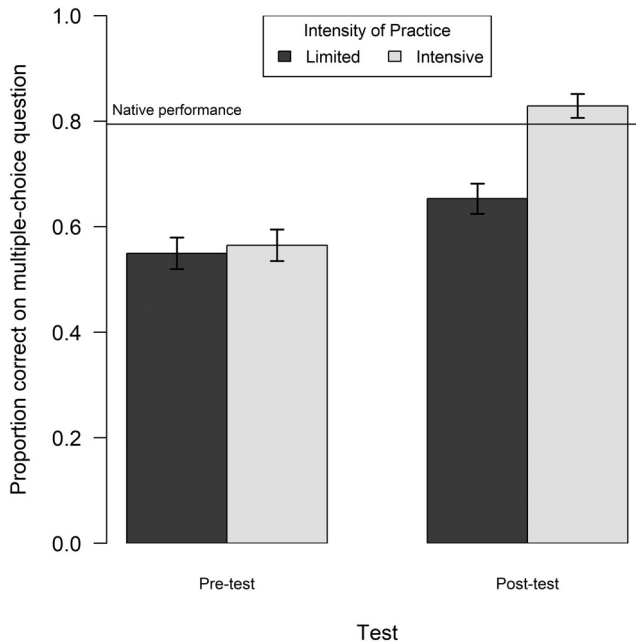
Figure 1. Mean proportion correct at pre-test and post-test for idioms that received limited and intensive practice. Horizontal line indicates mean native performance for the same idioms as taken from a previous study. The error bars represent SEs.

The fixed effects included in the final model were: (1) Test (reference category: Post-test), (2) Intensity of Practice (reference category: Limited), (3) Transparency, (4) Cross-Language Overlap (reference category: NE), (5) Set (covariate, reference category: A1), (6) LexTale score (covariate), (7) Test x Intensity of Practice, (8) Test x Cross-Language Overlap, (9) Test x Transparency, and (10) Test x Intensity of Practice x Transparency. Idioms (random intercept only) and Participants (random intercept and random slope of Test) were added to the model as random effects. The final model is presented in Table 4.

A significant interaction effect of Test and Intensity of Practice was observed ($\beta = -0.96$, $SE = 0.14$, $p < 0.001$). In the post-test, participants performed better on idioms that received intensive practice as opposed to idioms that received limited practice ($\beta = 1.03$, $SE = 0.11$, $p < 0.001$), whereas in the pre-test, idioms that received limited and intensive practice did not significantly differ (releveled version of the model: $\beta = 0.08$, $SE = 0.09$, $p > 0.05$).

In addition, the analysis revealed a significant interaction effect of Test and Cross-Language Overlap ($\beta = 0.36$, $SE = 0.20$, $p = 0.059$). In the pre-test no effects of Cross-Language Overlap were observed, whereas this was the case in the post-test. More specifically, in the post-test, participants performed worse on idiomatic expressions that do not exist in

Table 4. Final regression model with idiom knowledge as the dependent variable.

Fixed effects		Beta	Std. Error	z value
(Intercept)		0.8577	0.2399	3.575 ***
Pre-test vs. Post-test		−0.1925	0.2057	−0.936
Intensive vs. Limited		1.0321	0.1129	9.143 ***
Transparency		0.4765	0.1529	3.115 **
Cross-language overlap DW vs. NE		−0.4475	0.2202	−2.032 *
Cross-language overlap nW vs. NE		0.5802	0.2681	2.164 *
Cross-language overlap AW vs. NE		0.8703	0.3008	2.893 **
Set A2 vs. A1		−0.4759	0.1745	−2.728 **
Set B1 vs. A1		−0.3903	0.1766	−2.210 *
Set B2 vs. A1		−0.2742	0.1792	−1.530
LexTale score		0.3226	0.0637	5.065 ***
Pre-test x Intensive		−0.9569	0.1437	−6.657 ***
Pre-test x DW		0.2968	0.2067	1.435
Pre-test x nW		−0.8220	0.2502	−3.285 **
Pre-test x AW		−1.0838	0.2809	−3.858 ***
Intensive x Transparency		−0.4328	0.1540	−2.810 **
Pre-test x Transparency		−0.3433	0.1579	−2.174 *
Pre-test x Intensive x Transparency		0.3683	0.1952	1.886
Random effects		**Variance**	**Std. Deviation**	**Corr**
Idiom	Intercept	0.1536	0.3919	
Participant	Intercept	0.0856	0.2926	
	Test	0.1235	0.3514	0.13

Note.$*p < 0.05$ $**p < 0.01$ $***p < 0.001$.

their L1, German, (NE) as opposed to idioms that are word-to-word equivalents (AW, $\beta = 0.87$, $SE = 0.30$, $p < 0.01$) and idioms that have a number of content words in common (nW, $\beta = 0.58$, $SE = 0.27$, $p < 0.05$). In addition, in the post-test participants more often correctly recognized the meanings of idioms that do not exist in German, than meanings of idioms that do exist in German, but in totally different words (DW, $\beta = -0.45$, $SE = 0.22$, $p < 0.05$). A releveled version of the model showed no differences between idioms that have word-to-word equivalents in German and idioms that have a number of content words in common ($\beta = -0.29$, $SE = 0.23$, $p > 0.05$). See Figure 2 for a visual presentation of the interaction effect.

The analysis revealed a marginally significant three-way interaction effect with Test, Intensity of Practice and Transparency ($\beta = 0.36$, $SE = 0.20$, $p = 0.059$) (see Figure 3). By looking at the underlying two-way interactions, we gained more insight into the nature of the three-way interaction effect. In the post-test, the effect of Transparency appeared to be significantly smaller for the idioms that were intensively practiced than for the idioms that only received limited practice ($\beta = -0.43$, $SE = 0.15$, $p < 0.01$) (see right panel in Figure 3). More specifically, an effect of Transparency was absent for the intensively practiced idioms in the post-test (releveled version of the model: $\beta = 0.04$, $SE = 0.17$, $p > 0.05$), whereas a significant positive effect was observed for idioms that received limited practice ($\beta = 0.48$, $SE = 0.15$, $p < 0.01$). However, a releveled version of the model revealed no such differences
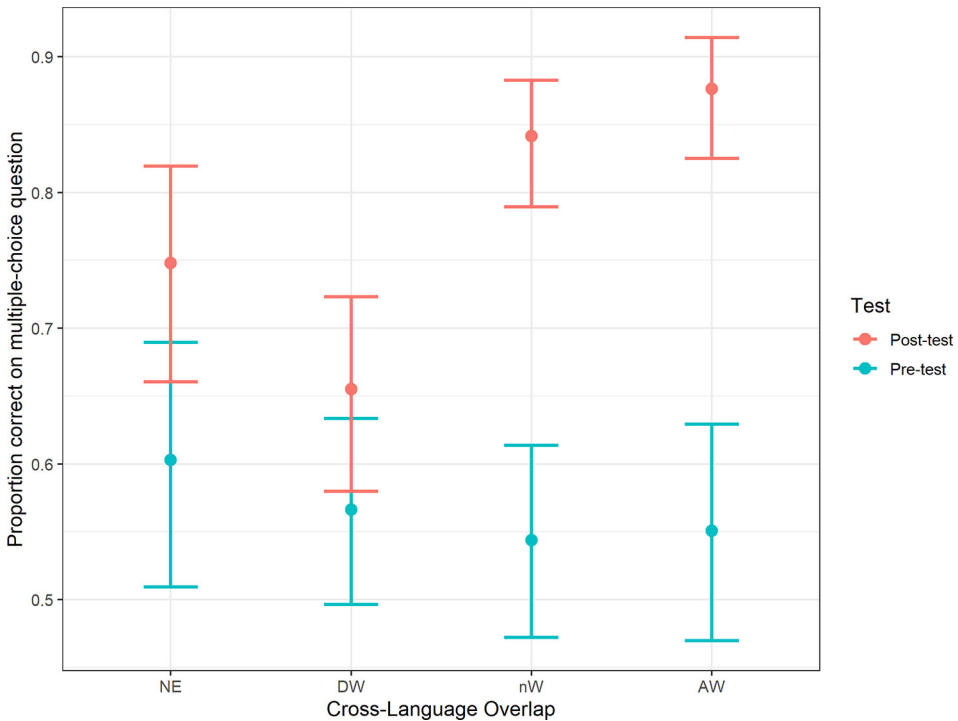
Figure 2. The interaction effect Test x Cross-Language Overlap. Effect of Cross-Language Overlap visible for the Post-test only. Error bars represent SEs.

in the pre-test ($\beta = -0.06$, $SE = 0.12$, $p > 0.05$). Here, an effect of Transparency was absent for both limited ($\beta = 0.13$, $SE = 0.14$, $p > 0.05$) and intensively practiced idioms ($\beta = 0.07$, $SE = 0.14$, $p > 0.05$).

Reading modality

Table 5 presents the mean proportions correct on the multiple choice question split by Test and Reading Modality only for a subset of the data (i.e., the idiomatic expressions that received intensive practice). See also Figure 4.

To investigate whether Reading Modality, either silent or aloud, has a differential effect on L2 idiom learning, the pre-test and post-test scores were submitted to logistic mixed effects regression analysis. Again, the performance on the multiple-choice question (correct/incorrect) was used as a dependent variable in the analysis.

The final model included the predictors (1) Test (reference category: Post-test), (2) Reading Modality (reference category: Silently), (3) LexTale score (covariate), and (4) Test x Reading Modality as fixed effects. Participants (random intercept and random slope of Test) and Idioms (random intercept only) were included as random effects (See Table 6).
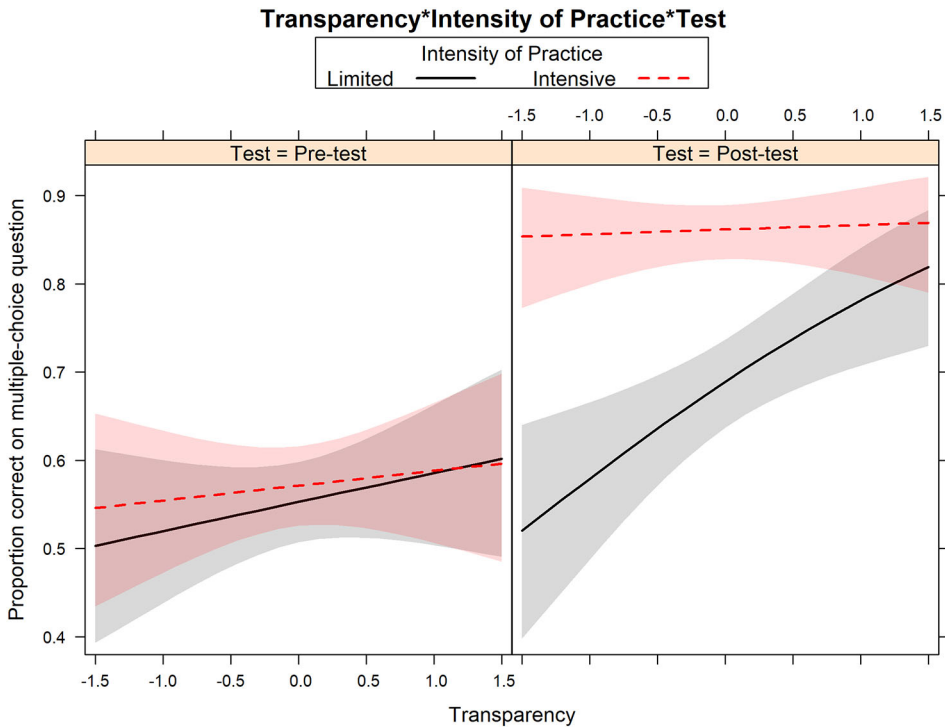
Figure 3. The interaction effect Test x Intensity of Practice x Transparency. Positive effect of Transparency visible only in the post-test (right panel) for the idioms that received limited practice. Error bands are based on SEs.

Table 5. Mean proportions correct and SDs on the multiple choice question.

	Reading Modality	
Test	Silent	Aloud
Pre-test	0.59 (0.49)	0.54 (0.50)
Post-test	0.85 (0.36)	0.81 (0.39).

The analysis revealed only significant effects for Test and LexTale score. Participants performed worse on the pre-test than on the post-test for both the idioms that were read aloud ($\beta = -1.56$, $SE = 0.19$, $p < 0.001$) and read silently ($\beta = -1.62$, $SE = 0.19$, $p < 0.001$). Vocabulary knowledge, as measured by LexTale, positively influenced performance on the multiple-choice question ($\beta = 0.04$, $SE = 0.009$, $p < 0.001$). No significant interaction effect was observed between Reading Modality and Test ($\beta = 0.06$, $SE = 0.21$, $p > 0.05$).

Discussion

In this paper, we employed a CALL system that provides automatic corrective feedback to investigate the effect of intensity of practice and
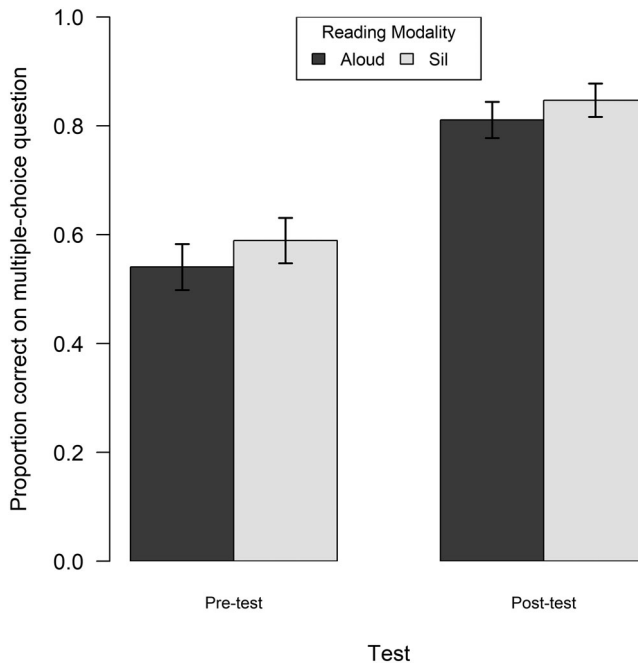
Figure 4. Mean proportion correct at pre-test and post-test for intensive practice idioms that were read aloud and silently. The error bars represent SEs.

Table 6. Final regression model with idiom knowledge as the dependent variable.

Fixed effects		Beta	Std. Error	z value
(Intercept)		−0.3609	0.6350	−0.568
Pre-test vs. Post-test		−1.6185	0.1938	−8.353 ***
Aloud vs. Silently		−0.2809	0.1687	−1.665
LexTale score		0.0369	0.0094	3.933 ***
Pre-test vs. Post-test x Aloud vs. Silently		0.0575	0.2108	0.273
Random effects		**Variance**	**Std. Deviation**	**Corr**
Idiom	Intercept	0.2018	0.4492	
Participant	Intercept	0.0985	0.3138	
	Test	0.4202	0.6482	0.33

Note. *$p < 0.05$ **$p < 0.01$ ***$p < 0.001$.

reading modality on learning Dutch L2 idiomatic expressions, as well as the impact of relevant idiom properties like degree of transparency and cross-language overlap.

The practice focused on meaning and form provided through our CALL system turns out to impact L2 idiom learning, and L2 learners could even achieve native-like performance. Even limited practice involving only two presentations of each idiomatic expression turned out to be effective. However, while intensive practice was effective for both opaque and transparent idioms, limited practice was effective for the more transparent idioms, but not for the more opaque ones. These findings are in line with our research hypotheses about the positive effect of practice and its relation to idiom transparency.

With respect to the effect of cross-language overlap, we found that L2 learners had more difficulties with idioms that do not exist in their L1, than with idioms that have word-to-word correspondents and idioms that have a number of content words in common. These findings align with our research hypothesis. However, the L2 learners more easily acquired idioms that do not exist in their L1, (Dutch: goed uit de verf komen = to make a good impression) than idioms that do exist in German, but in totally different words (Dutch: water naar de zee dragen = German: Eulen nach Athen tragen; English equivalent: to carry coals to Newcastle).

On the one hand, this seems to contrast with findings by Charteris-Black (2002) that idioms with equivalent concepts were easier to learn. On the other hand, it is in line with results by Irujo (1986), who found that only idioms that were identical or similar in form were easier to comprehend, whereas idioms that were different in form were more problematic. The present study adds to this line of research by indicating that idioms with no overlap in form are also more difficult to learn and require more intensive practice.

Surprisingly both transparency and cross-language overlap did not have a significant impact on idiom recognition at pre-test. In the post-test we found a significant effect of transparency in the limited practice condition, but not for intensive practice. Apparently, the participants in this study did not use this knowledge in the first place, but actually learned to exploit this information as a result of practice with the idiomatic expressions. The effect of transparency is visible in the limited practice condition where L2 learners have probably become aware of this property and need to use it to arrive at the meaning of L2 idioms. In the intensive practice condition they apparently received enough practice to learn the idioms without having to resort to transparency. Similar results with respect to cross-language overlap were obtained by Türker (2019), who also found no significant differences at pre-test between three types of idioms that were either (a) identical in form and meaning, (b) different in meaning, but identical in form or (c) existed in L2 only.

As to the manipulation of reading modality, we expected reading aloud to have a more positive effect on L2 idiom learning than reading silently, but this was not borne out by the results. We also expected this type of reinforcement activity to be more effective than the copying task investigated by Stengers et al. (2016), exactly for the reasons these authors mentioned. With reference to Levels of Processing Theory (Cermak & Craik, 1979; Craik & Lockhart, 1972), they argued that the copy-pasting activity was probably not engaging enough and therefore unlikely to leave persistent memory traces (Stengers et al., 2016, p. 296).

As a more elaborative alternative, they suggested copying by typing or by handwriting, as these would require more focus on the linguistic form than simple copy-pasting as in their study. Both our reading activities, silently and aloud, are tasks that draw attention to the orthographic form, with reading aloud requiring more phonological processing than silent reading. It is therefore possible that in this respect the two tasks were not sufficiently different from each other to bring about a difference in performance.

Another possible explanation, also suggested by Stengers et al. (2016) with respect to their findings, is that the idiom recognition test in the post-test was not completely appropriate to measure the effect of this kind of reinforcement practice. In order to check whether an effect of reading modality was observed in another type of exercise in the post-test, the sentence completion exercise, we examined these results in a post-hoc analysis. The same pattern of results was found for both reading modalities. Proportion correct for the idioms that had been practiced by reading aloud was 0.74, while the proportion correct was 0.77 for idioms that had been read silently. Finally, another possible explanation for the lack of a facilitatory effect of oral reading practice could be that the amount of oral practice was not sufficient within the context of our learning experiment.

Our results indicate that with sufficient exposure and practice, L2 learners demonstrate a level of meaning recognition equivalent to that of native speakers. This suggests that the underlying mechanisms of the two groups are similar. In fact, recent hybrid models of idiom processing in both L1 and L2 highlight the role of idiom properties, such as transparency and cross-language overlap, during idiom processing (Libben & Titone, 2008; Titone et al., 2015). These factors appear to exert a similar influence on idiom processing and on L2 idiom learning as observed in our study: Transparency is found to facilitate idiom processing in both native speakers (see Libben & Titone, 2008 and van Ginkel & Dijkstra, 2020) and L2 learners (Titone et al., 2015), and Cross-Language Overlap positively affected L2 idiom processing (Titone et al., 2015). Therefore, differences between the two groups seem to result from differences in exposure, rather than from differences in the underlying mechanisms involved in idiom processing.

Additional research is required to shed more light on the lack of a facilitatory effect of oral reading practice. In our study, the specific implementation of the reinforcement activity - reading aloud or silently - may not have been optimal or too short in duration to show any effect. As the current study only included comprehension activities in the pre- and post-test, it may be worthwhile for future research to include idiom

production activities as well, as these would give a more complete picture of L2 idiom learning. Another factor that could be included in future studies is linguistic distance. The L1 and L2 involved in this study are rather close, which is known to facilitate learning. It would be interesting to conduct similar experiments with an L1 and L2 that lie further apart, to see how transparency and cross-language overlap affect idiom learning in that case. Finally, in our future research we intend to investigate idiom learning after incorporating automatic speech recognition in the CALL system, which allows a more natural and possibly more rewarding form of interaction for the learner.

Conclusions

The present study allows us to conclude that CALL practice with a focus on meaning and form is effective for learning L2 idioms and that the degree of practice needed depends on the properties of the idioms. L2 learners can achieve or even exceed native-like performance. Higher intensity of practice leads to significantly better performance. Limited practice is effective for learning idioms that are transparent and idioms that contain similar words in L1. With intensive practice the differences in gains are smaller. Put otherwise, more practice is required for opaque idioms, for idioms that do not exist in L1, and for those that do exist in L1, but with a completely different form. Practicing reading idioms aloud does not lead to significantly higher performance than reading idioms silently.

The study has implications from both a theoretical and a teaching perspective. From a theoretical perspective, the findings indicate that differences between native speakers and L2 learners result from differences in exposure, rather than from differences in the underlying mechanisms involved in idiom acquisition. With respect to teaching practice, the study provides insights into which properties make L2 idioms more difficult to learn and how this knowledge could be exploited in idiom teaching. In teaching, specific attention should be paid to opaque idiomatic expressions, to idioms that do not exist in the learner's L1, and to idioms that do exist in the L1, but with a completely different form. Last but not least, the study presents an effective and ecologically valid way of facilitating L2 idiom practice, which is usually an unattended area in L2 classes, through a properly designed CALL system. L2 learners can learn from only two idiom presentations and with intensive practice they can reach native-like proficiency.

Note

1. Project website: https://isla.ruhosting.nl

Acknowledgments

Disclosure statement

Funding

Notes on contributors

Catia Cucchiarini is a Principal Investigator in the research group Language and Speech, Learning and Therapy and the Centre for Language and Speech Technology of the Radboud University Nijmegen, the Netherlands. Her research addresses speech processing, phonetics, second language acquisition, Computer Assisted Language Learning and the use of advanced language and speech technologies in this context.

Ferdy Hubers is a post-doc researcher at the Centre for Language Studies at Radboud University Nijmegen, the Netherlands. His research interests include idiomatic second language acquisition, Computer Assisted Language Learning, and neuro/psycholinguistics.

Helmer Strik is Associate Professor in Speech Sciences and Technology at the Centre for Language Studies and the Centre for Language and Speech Technology of the Radboud University Nijmegen, the Netherlands, and is co-founder and CSO of the spin-off company NovoLearning, which specializes in speech technology-based learning. He has conducted research in speech production, speech processing, and speech technology with particular focus on Automatic Speech Recognition (ASR) and its contribution to e-learning and e-health applications.

ORCID

Ferdy Hubers http://orcid.org/0000-0002-2298-6013

References

Ahmadi, A., Sahragard, R., & Babaie Shalmani, H. (2017). Anthropomorphism – matters or not? On agent modality and its implications for teaching English idioms and design decisions. *Computer Assisted Language Learning*, *30*(1–2), 149–172. doi:10.1080/09588221.2017.1284132

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi:10.18637/jss.v067.i01

Beck, S. D., & Weber, A. (2016). Bilingual and monolingual idiom processing is cut from the same cloth: The role of the L1 in literal and figurative meaning activation. *Frontiers in Psychology*, *7*, 1–16. doi:10.3389/fpsyg.2016.01350

Boers, F., Demecheleer, M., & Eyckmans, J. (2004). Cross-cultural variation as a variable in comprehending and remembering figurative idioms. *European Journal of English Studies*, *8*(3), 375–388. doi:10.1080/1382557042000277449

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*(3), 245–261. doi:10.1191/1362168806lr195oa

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*(1), 43–62. doi:10.1177/1362168806072460

Boers, F., & Lindstromberg, S. (2012). Experimental and Intervention Studies on Formulaic Sequences in a Second Language. *Annual Review of Applied Linguistics*, *32*, 83–110. doi:10.1017/S0267190512000050

Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. In F. Boers & S. Lindstromberg (Eds.), *Cognitive Linguistic approaches to teaching vocabulary and phraseology* (pp. 189–216). Berlin, Germany: Mouton De Gruyter.

Boers, F., Píriz, A. M. P., Stengers, H., & Eyckmans, J. (2009). Does pictorial elucidation foster recollection of idioms? *Language Teaching Research*, *13*(4), 367–382. doi:10.1177/1362168809341505

Bonin, P., Méot, A., & Bugaiska, A. (2013). Norms and comprehension times for 305 French idiomatic expressions. *Behavior Research Methods*, *45*(4), 1259–1271. doi:10.3758/s13428-013-0331-4

Caillies, S. (2009). Descriptions de 300 expressions idiomatiques: familiarité, connaissance de leur signification, plausibilité littérale, «décomposabilité» et «prédictibilité. *L'Année Psychologique*, *109*(3), 463–508. doi:10.4074/S0003503309003054

Cermak, L., & Craik, F. I. M. (1979). *Levels of processing in human memory* (L. Cermak & F. I. M. Craik, Eds.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Chan, T., & Liou, H.-C. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb – noun collocations. *Computer Assisted Language Learning*, *18*(3), 231–251. doi:10.1080/09588220500185769

Charteris-Black, J. (2002). Second language figurative proficiency: A comparative study of Malay and English. *Applied Linguistics*, *23*(1), 104–133. doi:10.1093/applin/23.1.104

Cieślicka, A. B. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, *22*(2), 115–144. doi:10.1191/0267658306sr263oa

Citron, F. M. M., Cacciari, C., Kucharski, M., Beck, L., Conrad, M., & Jacobs, A. M. (2016). When emotions are expressed figuratively: Psycholinguistic and Affective Norms of 619 Idioms for German (PANIG). *Behavior Research Methods*, *48*(1), 91–111. doi:10.3758/s13428-015-0581-4

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72–89. doi:10.1093/applin/amm022

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. doi:10.1016/S0022-5371(72)80001-X

Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, *28*(2), 130–144. doi:10.1080/09588221.2013.803982

Durrant, P., & Schmitt, N. (2010). Adult learners 'retention of collocations from exposure. *Second Language Research*, *26*(2), 163–188. doi:10.1177/0267658309349431

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, *42*(3), 375–396. doi:10.1002/j.1545-7249.2008.tb00137.x

Elshoff, J. (2014). Designing a task based curriculum for intensive language training. In S. Jager, L. Bradley, E. J. Meima, & S. Thouësny (Eds.), *CALL Design: Principles and Practice; Proceedings of the 2014 EUROCALL Conference* (pp. 66–71). Dublin: Search-publishing.net.. doi:10.14705/rpnet.2014.000196

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, *8*(15), 1–27. doi:10.18637/jss.v008.i15

Futagi, Y., Deane, P., Chodorow, M., & Tetreault, J. (2008). A computational approach to detecting collocation errors in the writing of non-native speakers of English. *Computer Assisted Language Learning*, *21*(4), 353–367. doi:10.1080/09588220802343561

Geluso, J. (2013). Phraseology and frequency of occurrence on the web: Native speakers' perceptions of Google-informed second language writing. *Computer Assisted Language Learning*, *26*(2), 144–157. doi:10.1080/09588221.2011.639786

Hsu, J. T., & Chiu, C. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *The Asian EFL Journal*, *10*(1), 181–204.

Hubers, F., Cucchiarini, C., Strik, H., & Dijkstra, T. (2019). Normative data of Dutch idiomatic expressions: Subjective judgments you can bank on. *Frontiers in Psychology*, *10*, 1–15. doi:10.3389/fpsyg.2019.01075

Hubers, F., van Ginkel, W., Cucchiarini, C., Strik, H., & Dijkstra, T. (2018). *Normative data on Dutch idiomatic expressions: Native speakers*. DANS. [Dataset]. doi:10.17026/dans-zjx-hnsk.

Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, *20*(2), 287–304. doi:10.2307/3586545

Irujo, S. (1993). Stearing clear: Avoidance in the production of idioms. *International Review of Applied Linguistics in Language Teaching*, *31*(3), 205–219.

Kecskes, I. (2007). Formulaic language in English Lingua Franca. In I. Kecskes & L. R. Horn (Eds.), *Explorations in pragmatics: Linguistic, cognitive and intercultural aspects* (pp. 191–219). Berlin and New York, NY: Mouton de Gruyter.

Kovecses, Z., & Szabco, P. (1996). Idioms: A view from cognitive semantics. *Applied Linguistics*, *17*(3), 326–255. doi:10.1093/applin/17.3.326

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). doi:10.18637/jss.v082.i13

Laufer, B. (2000). Avoidance of idioms in a second language: The effect of L1-L2 degree of similarity. *Studia Linguistica*, *54*(2), 186–196. doi:10.1111/1467-9582.00059

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. doi:10.3758/s13428-011-0146-0

Levy, M., Hubbard, P., Stockwell, G., & Colpaert, J. (2015). Research challenges in CALL. *Computer Assisted Language Learning*, *28*(1), 1–6. doi:10.1080/09588221.2014.987035

Libben, M. R., & Titone, D. A. (2008). The multidetermined nature of idiom processing. *Memory & Cognition*, *36*(6), 1103–1121. doi:10.3758/MC.36.6.1103

Liontas, J. (2002). Context and idiom understanding in second languages. *EUROSLA Yearbook*, *2*, 155–185. doi:10.1075/eurosla.2.11lio

Nesselhauf, N., & Tschichold, C. (2002). Collocations in CALL: An investigation of vocabulary-building software for EFL. *Computer Assisted Language Learning*, *15*(3), 251–279. doi:10.1076/call.15.3.251.8190

Neuner, G., KrüGer, M., &., & Grewer, U. (1981). *Übungstypologie zum kommunikativen Deutschunterricht*. Berlin: Langenscheidt.

Nordmann, E., Cleland, A. A., & Bull, R. (2014). Familiarity breeds dissent: Reliability analyses for British-English idioms on measures of familiarity, meaning, literality, and decomposability. *Acta Psychologica*, *149*, 87–95. doi:10.1016/j.actpsy.2014.03.009

Nordmann, E., & Jambazova, A. A. (2017). Normative data for idiomatic expressions. *Behavior Research Methods*, *49*(1), 198–215. doi:10.3758/s13428-016-0705-5

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). London: Longman.

Qualtrics. (2005). *Qualtrics*. Provo, UT. Retrieved from http://www.qualtrics.com

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rezaee, A. A., Marefat, H., & Saeedakhtar, A. (2015). Symmetrical and asymmetrical scaffolding of L2 collocations in the context of concordancing. *Computer Assisted Language Learning*, *28*(6), 532–549. doi:10.1080/09588221.2014.889712

Shei, C.-C. (2008). Discovering the hidden treasure on the Internet: Using Google to uncover the veil of phraseology. *Computer Assisted Language Learning*, *21*(1), 67–85. doi:10.1080/09588220701865516

Shei, C.-C., & Pain, H. (2000). An ESL writer's collocational aid. *Computer Assisted Language Learning*, *13*(2), 167–182. doi:10.1076/0958-8221(200004)13:2;1-D;FT167

Skoufaki, S. (2008). Investigating the source of idiom transparency intuitions. *Metaphor and Symbol*, *24*(1), 20–41. doi:10.1080/10926480802568448

Steinel, M. P., Hulstijn, J. H., & Steinel, W. (2007). Second language idiom learning in a paired-associate paradigm: Effects of direction of learning, direction of testing, idiom imageability, and idiom transparency. *Studies in Second Language Acquisition*, *29*(03), 449–484. doi:10.1017/S0272263107070271

Stengers, H., Deconinck, J., Boers, F., & Eyckmans, J. (2016). Does copying idioms promote their recall?. *Computer Assisted Language Learning*, *29*(2), 289–301. doi:10.1080/09588221.2014.937723

Tabossi, P., Arduino, L., & Fanari, R. (2011). Descriptive norms for 245 Italian idiomatic expressions. *Behavior Research Methods*, *43*(1), 110–123. doi:10.3758/s13428-010-0018-z

Titone, D., Columbus, G., Whitford, V., Mercier, J., & Libben, M. (2015). Contrasting bilingual and monolingual idiom processing. In R. R. Heredia & A. B. Cieślicka (Eds.), *bilingual figurative language processing* (pp. 171–207). New York, NY: Cambridge University Press.

Titone, D., & Libben, M. (2014). Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation. *The Mental Lexicon*, *9*(3), 473–496. 10.1075/ml.9.3.05tit.

Türker, E. (2019). Idiom acquisition by second language learners: The influence of cross-linguistic similarity and context. *The Language Learning Journal*, *47*(2), 133–112. doi:10.1080/09571736.2016.1221441

van Ginkel, W., & Dijkstra, T. (2020). The tug of war between an idiom's figurative and literal meanings: Evidence from native and bilingual speakers. *Bilingualism: Language and Cognition*, *23*(1), 131–147. doi:10.1017/S1366728918001219

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, *63*(1), 91–120. doi:10.1111/j.1467-9922.2012.00729.x

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, *21*(4), 463–489. doi:10.1093/applin/21.4.463

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press. 10.1017/CBO9780511519772.

Wu, S., Witten, I. H., & Franken, M. (2010). Utilizing lexical data from a Web-derived corpus to expand productive collocation knowledge. *ReCALL*, *22*(1), 83–102. doi:10.1017/S0958344009990218

Yorio, C. A. (1989). Idiomaticity as an indicator of second language proficiency. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss* (pp. 55–72). Cambridge: Cambridge University Press. 10.1017/CBO9780511611780.005.

Appendix

Table A1. Dutch idiomatic expressions included in the CALL-based learning experiment.

#	Idiomatic expression	Meaning
1	niet goed uit de verf komen	niet goed bij anderen overkomen
2	aan de grond zitten	in slechte omstandigheden verkeren
3	een vinger in de pap hebben	invloed hebben op iets
4	het op zijn heupen krijgen	plotseling fanatiek bezig gaan
5	hoog van de toren blazen	opscheppen
6	iemand iets in de maag splitsen	iemand iets dwingen te doen
7	iemand iets op de mouw spelden	iemand iets wijsmaken
8	iets onder de knie hebben	iets goed kunnen
9	iets soldaat maken	iets opeten
10	iets uit de doeken doen	iets uitleggen
11	lange tenen hebben	snel beledigd zijn
12	naast zijn schoenen lopen	zich arrogant gedragen
13	op de tocht staan	in een bedreigde positie komen
14	tegen de lamp lopen	betrapt worden
15	tegen het plafond zitten	niet meer kunnen bereiken
16	boter bij de vis	contant betalen
17	bij iemand een wit voetje halen	bij iemand in de gunst proberen te komen
18	de kat op het spek binden	iemand in verleiding brengen
19	een klein hartje hebben	gauw bang zijn
20	een slag om de arm houden	iets onder voorbehoud afspreken
21	een wassen neus	niet van belang
22	hek van de dam	geen belemmeringen meer hebben
23	iemand in de kaart spelen	iemand onbedoeld helpen
24	iemand om zeep helpen	iemand vermoorden
25	koek en ei zijn	goede vrienden zijn
26	met zijn neus in de boter vallen	in een gunstige situatie terechtkomen
27	op de fles gaan	failliet gaan
28	op een laag pitje staan	minder aandacht krijgen
29	op zijn strepen staan	zijn eigen mening aanhouden
30	veel voeten in de aarde hebben	veel moeite kosten
31	voor spek en bonen	zonder mee te tellen
32	aan een zijden draadje hangen	weinig kans op succes hebben
33	aan het roer staan	de leiding hebben
34	als paddenstoelen uit de grond schieten	snel en in grote massa tevoorschijn komen
35	de koe bij de hoorns vatten	een flinke klus aanpakken
36	de mouwen opstropen	aan het werk gaan
37	één lijn trekken	dezelfde aanpak gebruiken
38	ergens mee in de maag zitten	niet goed weten wat te doen met iets
39	ergens zijn neus voor optrekken	zich te goed vinden voor iets
40	het niet breed hebben	niet veel geld hebben
41	aan de tand voelen	iemand streng ondervragen
42	iemand de wet voorschrijven	bepalen wat iemand moet doen
43	iemand iets in de schoenen schuiven	iemand onterecht beschuldigen
44	iemand op de huid zitten	iemand continu, op een vervelende manier, controleren
45	iemand tegen de schenen schoppen	vervelend doen tegen iemand
46	iemand wakker schudden	iemand krachtig aan iets herinneren
47	iemand tot in de wolken verheffen	iemand uitbundig prijzen
48	iets op poten zetten	iets nieuws opstarten
49	iets uit de losse pols doen	iets zonder voorbereiding uitvoeren
50	iets uit je duim zuigen	iets verzinnen
51	iets van tafel vegen	iets zonder overleg afwijzen
52	met de rug tegen de muur staan	geen keuze hebben
53	op het puntje van de tong liggen	niet op een woord kunnen komen
54	op zijn achterste benen staan	je ergens tegen verzetten
55	op zijn centen zitten	gierig zijn
56	tussen twee vuren zitten	van twee kanten bedreigd worden
57	voor aap staan	zichzelf belachelijk maken
58	voor iemand door het vuur gaan	voor iemand alles overhebben
59	geen hart in het lijf hebben	geen medelijden kennen
60	het hoofd koel houden	rustig blijven

het op zijn heupen krijgen

plotseling fanatiek bezig gaan

resterende tijd:

Figure A1. Screenshot of the PAL paradigm exercise. The upper phrase is the Dutch idiomatic expression, and the phrase below is the corresponding meaning. Participants were instructed to carefully read the idiom-meaning pair. No explicit task was formulated.

Vraag 1   1/ 30
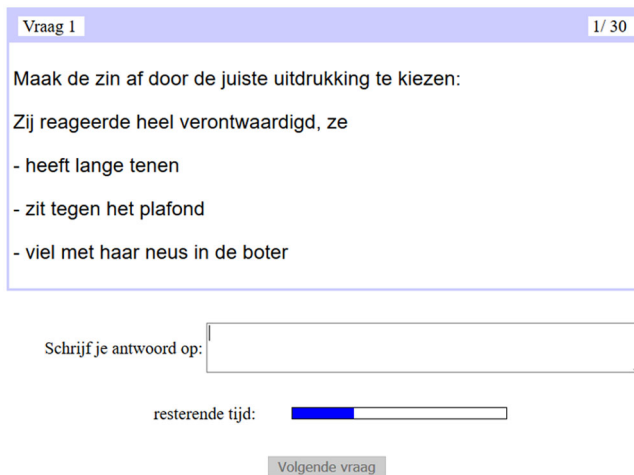
Vul het ontbrekende woord in:

Het verlegen meisje durfde uiteindelijk het verhaal uit de …
te doen

Schrijf je antwoord op:

resterende tijd:

Volgende vraag

Figure A2. Screenshot of the gap-fill exercise. Participants were instructed to type in the missing word. This word was always a noun that is part of an idiom. English translation from top to bottom: Question 1, Fill in the missing word, Write your answer, Remaining time, Next question.

Figure A3. Screenshot of the sentence completion exercise. Participants were instructed to type in the appropriate idiomatic expression based on the prompt. English translation from top to bottom: Question 1, Complete the sentence by choosing the right idiomatic expression, Prompt (sentence) + three idiomatic expressions, Write your answer, Remaining time, Next question.

Figure A4. Screenshot of the idiom selection exercise. Participants were instructed to type in the idiomatic expression that corresponded to the given meaning. English translation from top to bottom: Question 1, Choose the right idiomatic expression, Prompt (meaning of idiomatic expression) + three idiomatic expressions, Write your answer, Remaining time, Next question.

Figure A5. Screenshot of the corrective feedback participants received directly after providing the answer. The left panel shows the feedback in case of a correct answer and the right panel shows the feedback in case of an incorrect answer. English translation left panel: well done. English translation from top to bottom right panel: Incorrect, Correct answer, Input answer.

Figure A6. Screenshot of the reading aloud manipulation. Participants were instructed to read aloud the prompt as indicated by the record button. English translation from top to bottom: Question 1, Sentence + idiomatic expression, Record your answer, Remaining time, Next question.

Figure A7. Screenshot of the reading silently manipulation. Participants were instructed silently read the prompt as indicated by the absence of the record button. English translation from top to bottom: Question 1, Sentence + idiomatic expression, Remaining time, Next question.