

DATA-DRIVEN LEARNING OF COLLOCATIONS: LEARNER PERFORMANCE, PROFICIENCY, AND PERCEPTIONS

Nina Vyatkina, University of Kansas

This study explores the effects of Data-Driven Learning (DDL) of German lexico-grammatical constructions (verb-preposition collocations) by North-American college students with intermediate foreign language proficiency. The study compares the effects of computer-based and paper-based DDL activities as evidenced in learners' immediate and delayed performance gains, and explores changes in learners' proficiency and DDL perceptions as well as the influence of these factors on performance. The results show that both DDL types were equally effective for all learners, independent of their proficiency and perceptions, although gains measured by a more controlled production test (gap-filling) were superior to and longer lasting than gains measured by a less controlled production test (sentence-writing). Furthermore, immediate performance gains on different tasks were differently affected by learner proficiency and perceptions, while delayed gains showed no such effects. Finally, the study found that overall learner proficiency increased and that DDL was well received by learners and they expressed an intention to use it for independent learning in the future. This study fills gaps existing in DDL research by focusing on a second language other than English, comparing different DDL types, measuring delayed learning gains, and combining different outcomes measures in a multilevel modeling design.

Language(s) Learned in Current Study: German

Keywords: Computer-assisted Language Learning, Corpus, Data-driven Learning, Language Teaching Methodology, Learners' Attitudes

APA Citation: Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning & Technology*, 20(3), 159–179. Retrieved from <http://lt.msu.edu/issues/october2016/vyatkina.pdf>

Received: March 5, 2015; **Accepted:** September 28, 2015; **Published:** October 1, 2016

Copyright: © Nina Vyatkina

INTRODUCTION

Corpora, or large digital textual databases, have been attracting attention of language educators since their emergence in the 1960s. Corpus-based language teaching applications can be indirect or direct (Leech, 1997). In the indirect approach, teachers and learners use published corpus-informed pedagogical materials such as grammars, textbooks, and dictionaries, whereas in the direct approach, they actively interact with corpora with the purpose to discover and analyze patterns of language use (Römer, 2011). Such direct corpus-based applications have been singled out as a specific language learning and teaching method by Johns (1990), who dubbed it Data-Driven Learning (DDL). Research exploring DDL effects has been growing recently (see collections by Boulton & Pérez-Paredes, 2014; Frankenberg-Garcia, Flowerdew, & Aston, 2011; Leńko-Szymańska & Boulton, 2015; Thomas & Boulton, 2012), although this teaching method is still far from becoming common pedagogical practice (Römer, 2011).

Although research on DDL has demonstrated its significant benefits for language learning (for a meta-analysis, see Cobb & Boulton, 2015), there are still considerable gaps to be filled. One is that the bulk of available studies have targeted English as a Second or Foreign Language (ESL or EFL), whereas applications to other languages are few and far between. Another gap is that DDL has mostly been compared to non-DDL instruction methods, while the relative effects of different DDL types have been

underexplored. Most studies have also investigated exclusively either learner performance outcomes or learner attitudes. Studies that integrated multiple measures are rare. Furthermore, few studies explored delayed DDL effects.

This study aims to fill these gaps by exploring how the productive knowledge of German verb-preposition collocations by North-American college students was affected by the DDL type (computer-based and paper-based), time (pre-test, immediate post-test, and delayed post-test), task (gap-filling and sentence-writing), as well as learners' proficiency and DDL perceptions.

LITERATURE REVIEW

Theoretical Background

This study is grounded in usage-based language acquisition theories (e.g., Bybee, 2006; Robinson & Ellis, 2008) that conceive of grammar as an abstract cognitive representation of a language that is gradually constructed in the learner's mind from exposure to particular exemplars in the environment. One of the most important principles undergirding the usage-based approach adopted by corpus linguists and, more recently, some DDL researchers is inseparability of lexis and grammar (Chambers, 2005). The main units of analysis in this research are linguistic constructions that can consist of concrete items (e.g., words), abstract items (e.g., word classes), and combinations thereof (Ellis, 2014). One example of such a construction is collocation: "the characteristic co-occurrence of patterns of words" (McEnery, Xiao, & Tono, 2006, p. 345), such as adjective-noun or verb-preposition collocations. Collocations are an important aspect of the depth of second language (L2) knowledge (Barfield, 2013) but errors in using them persist even at advanced levels of L2 proficiency (Paquot & Granger, 2012). This issue has been addressed in DDL teaching practices and associated pedagogical principles, notably the noticing hypothesis and discovery learning (Flowerdew, 2015).

The noticing hypothesis proposes that increased noticing and awareness are beneficial for adult L2 learning (Robinson, 1995; Schmidt, 1990, 2001). Noticing can be facilitated by the use of two teaching techniques: input enrichment, or repeated exposure of learners to the target structure over a period of time, and input enhancement, or emphasizing the target structure by typographical means such as bolding and color marking (Sharwood Smith, 1993). Non-DDL research has shown that input enrichment and input enhancement can indeed facilitate the learning of collocations, at least their recognition and recall (Sonbul & Schmitt, 2013; Szudarski & Carter, 2016). In DDL, input enrichment is realized as corpora provide teachers and learners access to a large number of target structures in attested language use samples, which are hard to come by in a traditional language classroom, especially in foreign language learning settings. Input enhancement is realized through the use of concordancers, tools for automatic retrieval of collocations, some of which are integrated with corpora and others which exist as external tools. Concordancers supply search results as stacked lines with the search words highlighted in the middle (see [Appendix D](#)) and thus enhance the visibility of collocational patterns.

The second principle is inductive (discovery) learning (Bernardini, 2002), also associated with learner autonomy (Gavioli, 2009). Early DDL research primarily explored the strong version of DDL, in which learners perused corpora directly to discover patterns of language use (with some assistance from the teacher). Although there have been several success stories involving more advanced learners, lower proficiency learners often struggled with this approach and felt overwhelmed while working on (new) L2 material with a new method and a new technology (Boulton, 2010; Gavioli, 2005). These findings have prompted DDL researchers to create and explore modifications of DDL instructional designs that "can be plotted on a cline of learner autonomy, ranging from teacher-led and relatively closed concordance-based activities to entirely learner-centered corpus-browsing projects" (Mukherjee, 2006, p. 12). These DDL approaches have been termed *soft* and *hard* (Gabrielatos, 2005) or *hands-off* and *hands-on*, respectively (Boulton, 2012). Boulton (2010, 2012), in particular, has advocated the hands-off approach in initial

learner encounters with corpora at lower L2 proficiency levels and in institutional contexts where hands-on approaches are not feasible. The research body on the effectiveness of both hands-on and hands-off DDL has been growing and is reviewed in the next section.

Empirical Research on Hands-on and Hands-off DDL

Empirical DDL research has explored whether this method of instruction is effective (i.e., whether it results in learning gains), and efficient (i.e., whether it is better than non-DDL, or mostly deductive, methods). Cobb and Boulton (2015) conducted the first meta-analysis of quantitative DDL studies and showed that, although individual study effects tend to be small, the overall effect size for both effectiveness and efficiency was significant and high. As regards language foci, some studies employed broad improvement measures such as accuracy gains or the holistic quality of learner production. Other studies explored more specific lexico-grammatical foci and used short comprehension and production activities as data collection instruments (e.g., multiple-choice and gap-filling items). Studies of this latter type that focused on DDL of L2 collocations are most relevant to this study and are reviewed below. The overwhelming majority of these studies have been conducted in ESL or EFL instructional contexts and had university students at intermediate to advanced L2 proficiency levels as participants, which is typical of DDL research in general.

For hands-on interventions, significant advantages have been found for learning L2 English collocations: noun-verb collocations by first language (L1) Chinese speakers (Chan & Liou, 2005), verb-adverb collocations by L1 Macedonian speakers (Daskalovska, 2015), and a variety of collocations by L1 Arabic speakers (Cobb, 1997). Studies focusing on hands-off DDL interventions for teaching a variety of English collocations to students with different L1s (Arabic, Portuguese, French, Thai, Chinese) found them either significantly more efficient than traditional teaching methods (Frankenberg-Garcia, 2014; Koosha & Jafarpour, 2006) or equally effective with marginally higher DDL gains (Boulton, 2010; Scripicharn, 2003; Tian, 2005). In a rare study that focused on L2 German, Vyatkina (2016) found that hands-off DDL was more effective than traditional instruction for learning new verb-preposition collocations by low-intermediate L1 English learners, but both methods were equally effective for improving knowledge of previously learned collocations. Additionally, several of the abovementioned studies collected retrospective learner attitude data and reported on a generally positive perception of DDL activities by the learners.

What remains underexplored is the efficiency of specific DDL types compared to each other. Only a few studies focused on collocations have addressed this issue. Sun and Wang (2003) showed that inductive hands-on DDL activities led to greater gains in Chinese learners' knowledge of easier English collocations than deductive hands-on DDL activities, but there was no difference between DDL types for more complex collocations. Frankenberg-Garcia (2014) found that hands-off DDL with several concordance lines was not only better than traditional instruction with a dictionary but also better than hands-off DDL with a single concordance line for teaching English collocations to L1 Portuguese learners. These two studies are significant in that they have singled out specific DDL principles that make this type of instruction beneficial: inductive discovery and input enrichment.

Another research gap is the lack of studies comparing the outcomes of hands-on and hands-off DDL interventions, with only two recent studies tackling this issue. An exploratory case study by Yoon and Jo (2014) showed hands-off DDL to be more beneficial for the overall rate of self-correction of writing errors by Korean learners of English than hands-on DDL, although the learners liked the hands-on method better. The only quantitative study is Boulton's (2012), who compared a hands-on and a hands-off DDL intervention teaching English verb-infinitive and verb-subjunctive collocations to intermediate college-level EFL learners with L1 French. Notably, Boulton also explored correlations of the production outcomes with learners' proficiency and perceptions. The results showed that (a) hands-on and hands-off DDL were equally effective, (b) only hands-off performance outcomes correlated with L2 proficiency,

and (c) the correlation of learner production with perception was positive but not significant. The present study aims to continue this line of research by comparing learner performance outcomes following hands-on and hands-off DDL and integrating proficiency and perception outcomes as well as adding the task factor and a delayed performance test. In order to systematically account for interactions among all these factors, the study employs a multilevel modeling research design. The rationale for choosing verb-preposition collocations as the target construction is discussed in the next section.

The Target Feature

Prepositions, which are among the most frequent words in many languages, have been repeatedly pointed out as an area of difficulty for language learners. De Felice and Pulman (2009), for example, found that 12% of all errors in a learner English corpus were preposition errors. Kennedy and Miceli (2001) called this problem “the fatal lure of prepositions” and exemplified it by the way L1 English learners of Italian treated prepositions in their L2:

The students’ attention was often attracted to a preposition itself rather than to the words around it, on which it depended. In some cases, they treated a preposition as having a meaning in isolation, or as being in one-to-one correspondence with an English preposition. (p. 83)

This preferred learner approach is especially faulty in the case of prepositions that are bound to specific verbs, nouns, or adjectives. In these constructions, prepositions are closely tied to the preceding lexical word, invariable, and semantically empty (Biber, Johansson, Leech, Conrad, & Finegan, 1999). German, like English and many other languages, has verbs that subcategorize certain prepositional arguments. However, form-meaning mappings among languages often diverge, which causes transfer errors in learner language. For example, the German equivalent of the English *to wait for* is *warten auf* (*to wait on*), although the prototypical translation of the preposition *for* is *für*. In a learner corpus study, Nesselhauf (2003) demonstrated a strong L1 transfer effect in the use of English collocations by L1 German learners, including noun-preposition and verb-preposition collocations.

This difficulty is further exacerbated in the case of German, an inflectional language; noun phrases and personal pronouns that follow verb-preposition collocations must carry inflectional markers. For example, in the sentence *Ich warte auf den Zug* (*I am waiting for the train*), the inflected form of the definite article *den* indicates the masculine gender, accusative case, and singular number. Furthermore, whereas the noun gender belongs to the lexicon (e.g., *Zug* [*train*] is always masculine) and the number assignment depends on the context, the case (usually accusative or dative, and sometimes genitive) is assigned either by the verb or by the preposition in each specific collocation. General SLA research shows that L2 German learners acquire case in such constructions fairly late, after years of instruction (Baten, 2011). This study investigates whether a new method, DDL, may help tackle this complex lexico-grammatical L2 phenomenon that has been refractory to traditional teaching methods.

DESIGN

Research Questions

The following research questions (RQs) were explored:

1. Did learner written performance improve following DDL instruction and were the gains retained a month later? Was performance affected by task, proficiency, and perception?
2. Which DDL method was more efficient: hands-on or hands-off? Was this effect modulated by time, task, proficiency, and perception?
3. Did learners’ proficiency and DDL perception change over time, and what was the relationship between proficiency, perception, and time?

Participants and Institutional Setting

The DDL intervention was incorporated into a German course taught by the researcher at a large public North American university. 11 students were enrolled in the course, however, the study reports only on 10 participants (five male, five female) who were present on all DDL test days. The mean age of the participants was 21 (ranged 18–24), and all of them had American English as their L1. Nine participants were junior and senior university students with German as their major or minor, and one participant was a High School student who was taking the course for college credit. All students but one had visited German-speaking countries for times ranging from several weeks to several years. The duration of participants' previous German study and their exposure to German varied, but on average, their German proficiency was at the B1 level according to the Common European Framework of Reference (CEFR; Council of Europe, 2001). Finally, one participant had had some previous knowledge of language corpora, whereas the other nine participants had none.

The course was titled Advanced German I, and the class met two times a week for 75 mins over 16 semester weeks (SWs). The course was oriented toward development of general proficiency in German and combined extensive reading, writing, discussion, and grammatical analysis activities. Beyond that, the course also aimed to develop corpus literacy (Mukherjee, 2006) by including regular hands-on and hands-off corpus assignments (similar to Boulton, 2012) constituting 30% of the total course grade.

Data Collection Timeline

The data collection timeline is presented in Table 1. During the first and last lab sessions (SW1 and 15), participants filled out a DDL receptivity questionnaire and took a standardized German proficiency test. They also filled out a language background questionnaire in SW1. For the DDL experiment that supplied performance data for this study, all learners participated first in the hands-on condition (SW 11) and, during the next class 5 days later, in the hands-off condition (SW 12). Each experiment took one 75-minute-long class period. Participants took the pre-test (ca. 10 min.) at the beginning and the post-test (ca. 10 min.) at the end of class, and the 55 minutes in between were spent for the DDL interventions. In SW 16, the delayed post-test was administered. There was no explicit instructional focus on the target items between the immediate and delayed post-test.

Table 1. Data Collection Timeline

Aug. 28 (SW1)	Nov. 6 (SW11)	Nov. 11 (SW12)	Dec. 4 (SW15)	Dec. 9 (SW16)
Proficiency and perception pre-test; language background questionnaire	Performance hands-on: pre-test, treatment, and post-test	Performance hands-off: pre-test, treatment, and post-test	Proficiency and perception post-test	Performance: delayed post-test

DDL Instruction Procedures

The corpus used in the course was part of the [Digital Dictionary of German \(DWDS\)](#). DWDS is a large open access digital collection that includes, among many other corpora and corpus tools, the core corpus (*Kernkorpus*) of 20th century German texts and a concordancer. All participants experienced DDL both on computer (hands-on) and on paper (hands-off) to approximately the same extent over the semester. More specifically, students completed hands-on activities during six biweekly computer lab sessions, each taking the whole class period. Hands-off activities were based on teacher-prepared worksheets used during regular class meetings. Both types of DDL work were also occasionally assigned as homework. All these activities provided the learners with ample exposure to DDL methods before the experiment, following Boulton's (2012) approach:

The activities themselves were typical of those discussed in DDL research, involving induction from authentic concordances, sorting and interpreting data, testing and matching rules, and so on. The activities allowed learners to experience a range of typical DDL tasks: amongst other things, printed concordances featured matching and gap-fill exercises, while computer searches became rather freer over the course. (p. 155)

20 verb-preposition collocations¹ (Appendix A) were selected for the study, all of them appearing in the text of the novel that students read and discussed as the main course text (Moers, 2002) in conformity with the principle of relevance shown to promote learning (Boulton, 2010; Hulstijn & Laufer, 2001). Furthermore, these collocations appear in common German teaching materials for intermediate to advanced levels. All of the focal collocations had a frequency of occurrence of at least 45 per million words in the DWDS core corpus. Each DDL module (hands-on and hands-off) focused on one half of these collocations matched by their corpus frequency.

Each test sheet contained sentences (Appendix B) with focal collocations from the course novel. The pre-test simultaneously served as the first part of the instructional intervention. It functioned as an awareness-raising exercise as learners realized gaps in their knowledge of the focal collocations (despite repeated exposure during years of learning German and a recent encounter in the text of the novel). Both teaching modules followed the DDL 3 Is (*Illustration-Interaction-Induction*) principle (Carter & McCarthy, 1995) and also included the fourth *Intervention* step (Flowerdew, 2009). Following the pre-test, participants received worksheets with instructions to work with DWDS concordances of the focal verbs following a model (*Illustration*). Specifically, they were asked to underline the prepositional phrase following the verb, decide about the case, and to write out the verb-preposition-case collocation following a model (e.g., *warten auf* + accusative). Then, they were asked to compare their results with a partner (*Interaction*) and to come to a joint solution (*Induction*). The instructor then checked all results to ensure they were accurate (*Intervention*). The difference between the conditions was the following: In the hands-on condition, learners searched the corpus for the focal verbs, read concordances, copied two examples for each verb, and pasted them in their worksheets for analysis (Appendix C). In the hands-off condition, the worksheets supplied four to seven concordance lines for each item with highlighted verbs and prepositions (Appendix D).

Data Collection Instruments and Scoring

Performance

The testing materials for performance outcomes for each condition (hands-on and hands-off) included a gap-filling task and a sentence-writing task. Each pre-test and immediate post-test included 10 gap-filling items and five sentence-writing items with five pre-assigned verbs randomly selected from the gap items. On the pre-tests, sentences with bolded focal verbs and gaps for focal prepositions were arranged in the sequence they appeared in the novel, and on the post-tests, they were scrambled (see Daskalovska, 2015). The delayed post-test included all 20 items from both treatment conditions with sentences following the order they appeared in the novel and all 10 verbs used in previous tests for sentence writing in a random order. In the gap-filling part, each accurately supplied preposition earned one point. In the sentence-writing part, each accurately supplied preposition earned one point and each accurately supplied inflection after each accurately supplied preposition earned one point. Therefore, the maximum number of points each learner could earn was 10 per test, per condition, per task.

Proficiency

Participant L2 proficiency was measured twice, in SW1 and SW16, during the regularly scheduled lab hour, with an official [online diagnostic test](#) administered by the onDaF Institute in Bochum, Germany, and proctored at the researcher's institution. Upon completion of this 40-minute-long cloze test, participants earned a certain number of points equivalent to the number of correctly filled gaps and were

automatically placed within CEFR bands (Eckes & Grotjahn, 2006). The descriptive statistics for the proficiency outcomes are presented in Table 2. It shows that prior to the course, the average L2 proficiency was slightly below the onDaF core range values for the B1 CEFR level (75–80), whereas after the course, it was slightly above that core range. Only post-course proficiency outcomes (available for all 10 participants) were used to answer RQ1 and RQ2, whereas temporal proficiency changes were addressed in answering RQ3 using proficiency outcomes for the nine students who completed both the pre-test and post-test.

Table 2. *Descriptive Statistics for Proficiency Outcomes*

onDaF Test	N	M	Min	Max	SD	N in CEFR bands		
						A2	B1	B2
Pre-course	9	73.78	36	115	26.88	3	4	2
Post-course	10	80.70	48	118	22.94	1	6	3

Perception

DDL perception data were collected by means of a written pre-course and a post-course receptivity questionnaire that partially replicated Boulton's (2012, p. 161) questionnaire and in which learners rated their expectations and satisfaction regarding DDL activities on a 5-point scale (Table 3). Only post-course perception outcomes are used to answer RQs 1 and 2, whereas temporal changes are addressed in answering RQ3. Additionally, as part of a cumulative final course assignment, students wrote open-ended commentaries about their experiences with DDL.

Table 3. *Descriptive Statistics for Perception (Receptivity) Outcomes*

Receptivity to Corpus Use		Pre-course		Post-course	
		M	SD	M	SD
1	The German corpus (DWDS) will be / has been easy to use for language learning purposes.	3.7	0.82	3.7	0.82
2	The German corpus (DWDS) will be / has been useful for learning German.	3.6	0.96	3.9	1.10
3	Corpus work will be / has been interesting.	3.4	1.08	4.1	1.20
4	I liked doing corpus activities on computer.	N/A	N/A	3.5	1.43
5	I liked doing activities with corpus concordances on paper.	N/A	N/A	3.8	1.14
Overall mean		3.57	0.13	3.8	0.22

Note. Questions adapted in part from Boulton (2012, p. 161).

RESULTS

Research Question 1

Did learner written performance improve following DDL instruction and were the gains retained a month later? Was performance affected by task, proficiency, and perception?

First, the pre-test and post-test performance scores of each of the 10 participants were inspected using simple counting. This analysis showed that nine of them improved their performance on the immediate post-test. Only one learner scored 15 points on the immediate post-test in comparison with 16 on the pre-

test in the hands-on condition, and the same learner scored 11 points on both the pre-test and immediate post-test in the hands-off condition. Further inspection showed that this lack of improvement was due to her performance only on sentence-writing items, whereas she improved on the gap-filling items. Furthermore, another learner scored 2 points lower on the immediate post-test than on the pre-test on sentence-writing items. Finally, all learners scored lower on the delayed post-test than on the immediate post-test but higher than on the pre-test, with only the same two learners who lacked improvement on the immediate post-test falling slightly below their pre-test scores on the delayed post-test.

Next, learner performance outcomes were explored statistically using multilevel regression models, which have been recently acknowledged as a preferred method in SLA research in comparison with more traditional ANOVA methods (Cunnings, 2012). The descriptive statistics are summarized in Table 4 and pre- to post-test changes are illustrated in Figure 1. One can see that the average pre-test knowledge was lower than 50% (out of 10 possible points) per each test task. Following the DDL modules, all outcomes increased on the immediate post-test and decreased on the delayed post-test, although not to the level of the pre-test.

Table 4. Descriptive Statistics for Performance Outcomes

Condition	Task	Pre-test		Post-test		Delayed Post-test	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
hands-on	gap	4.40	2.76	8.00	1.33	6.00	2.16
	sentence	4.20	2.97	6.90	2.33	4.70	2.58
hands-off	gap	3.50	1.96	7.60	1.65	4.70	2.41
	sentence	2.60	2.17	4.70	2.91	3.20	2.57

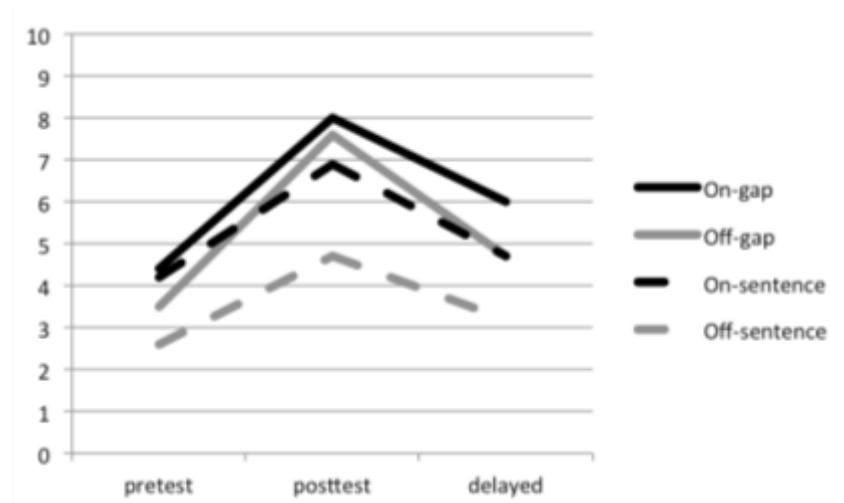


Figure 1. Average raw test scores by time, treatment method, and task.

Choosing the Method and Fitting the Model

To explore whether these changes were significant and how they were affected by task, learner proficiency, and learner perception, multilevel Poisson regressions were applied to the raw test scores treating them as counts and using z -tests associated with each parameter and test. This method was chosen because the raw test scores were not normally distributed. The values for the performance scores (dependent variable) at repeated measurements (Level-1) were nested within individual students (Level-

2), and the values for task, proficiency, and perception were factored in as predictor variables in addition to time. Because zero was outside the range of observed values for both proficiency and perception, they were both centered at their grand means to help interpret the other parameters. Post-test proficiency values were divided by 10 so that their scale was more similar to other variables, which helped the models converge on a solution.

In the initial model, proficiency and perception were both allowed to interact with time of measurement, which significantly improved the model with only time as a predictor of gap-filling scores ($\chi^2(6) = 22.26$, $p = .001$). But adding the interaction between proficiency and perception did not significantly improve the model for gap-filling scores ($\chi^2(1) = 1.18$, $p = .28$), and attempting to add any higher-order terms (i.e., three-way interactions between time, proficiency, and perception) prevented the model from converging on a solution due to lack of identification (probably because there were too few observations to support such a complex model). So the final model included proficiency, perception, two dummy codes for time, and the interaction of those dummy codes with proficiency and with perception. For the sentence-writing outcome, allowing proficiency and perception to interact with time only marginally improved the model ($\chi^2(6) = 12.58$, $p = .0502$), but the interaction terms themselves were not significant (all p -values > 0.10). Again, allowing proficiency and perception to interact with each other did not improve the model ($\chi^2(1) = 0.84$, $p = .36$), so no higher-order terms were included in the final model.

Immediate Post-test Scores

For students with average proficiency and perception, the immediate post-test gap-filling scores were significantly (on average, 2.24 times) higher than pre-test scores (95% CI [1.68, 3.03], $z = 5.34$, $p < .0001$). This average improvement was significantly greater for students with higher perception ($z = 2.45$, $p = .01$), such that each 1-unit increase in perception increased the average post- to pre-test ratio by 56% (95% CI [10%, 123%]). However, the average improvement was significantly less for students with higher proficiency ($z = -3.16$, $p = .002$), such that each 10-unit increase in proficiency decreased the average post- to pre-test ratio by 22% (95% CI = [9%, 33%]).

Holding proficiency and perception constant (at the mean or any other level, since the interaction was not significant), the immediate post-test sentence-writing scores were significantly (on average, 1.71 times) higher than pre-test scores (95% CI [1.27, 2.31], $z = 3.51$, $p = .0005$). Perception did not have a significant effect on sentence-writing scores ($z = -1.59$, $p = .11$) but proficiency did ($z = 3.69$, $p = .0002$), such that each 10-unit increase in proficiency resulted in 21.6% higher average scores (95% CI [9.3%, 38%]), controlling for time and perception.

Delayed Post-test Scores

Ignoring pre-test gap-filling scores, time (immediate vs. delayed post-test) did not significantly interact with perception ($z = -1.12$, $p = .26$), nor with proficiency ($z = 1.94$, $p = .052$). Controlling for proficiency and perception, delayed post-test gap-filling scores were significantly (on average, 66%) lower than immediate post-test scores (95% CI [51.2%, 84.8%], $z = -3.23$, $p = .001$), but delayed post-test scores were still significantly (on average, 48%) higher than pre-test scores (95% CI [8%, 105%], $z = 2.41$, $p = .02$). Controlling for proficiency and perception, delayed post-test sentence-writing scores were significantly (on average, 68%) lower than immediate post-test scores (95% CI [51%, 90%], $z = -2.64$, $p = .008$), and they were not significantly different from pre-test scores ($z = 0.91$, $p = .36$).

Summary of RQ1 Results

DDL was found to be effective for short-term performance gains on both the gap-filling and sentence-writing task. Furthermore, the higher DDL perceptions learners had, the more they learned on the gap task, whereas perception had no effect on the performance on the sentence task. Additionally, less-proficient learners improved more on the gap task than more-proficient learners, but the pattern was the opposite for the sentence task. Furthermore, all performance outcomes declined a month later in

comparison with the immediate post-test. However, the delayed gap outcomes still remained higher than the pre-test outcomes, whereas the sentence outcomes were not significantly different from the pre-test outcomes. Finally, learners' proficiency and perception did not have an effect on their delayed post-test outcomes.

Research Question 2

Which DDL method was more efficient: hands-on or hands-off? Was this effect modulated by time, task, proficiency, and perception?

Choosing the Method and Fitting the Model

To explore this question, test gains between the pre-test and the immediate post-test as well as between the pre-test and the delayed post-test were compared across the treatment methods and tasks. Because test gains (as opposed to raw test scores) were normally distributed, a multilevel linear regression method was used to explore significance of these differences with *t*-tests. Starting with a model that included the method (hands-on or hands-off), task (gap-filling or sentence-writing), and their interaction to predict *immediate* gains, adding proficiency and perception (and all interactions) did not significantly improve the fit of the model ($\chi^2(12) = 19.02, p = .09$). So proficiency and perception were left out of the model.

Starting with a model that included the method (hands-on or hands-off), task (gap-filling or sentence-writing), and their interaction to predict *delayed* gains, adding proficiency and perception (and all interactions) significantly improved the fit of the model ($\chi^2(12) = 22.35, p = .03$), but inspection of the individual effects revealed the 4-way interaction was not significant ($t(22) = -0.25, p = .81$), so it was dropped from the model. The model with all 2- and 3-way interactions still fit significantly better than the original model ($\chi^2(11) = 22.29, p = .02$), but inspection of the individual effects revealed no significant 3-way interactions (all *p*-values > .064), so they were dropped from the model. The model with all 2-way interactions did not fit significantly better than the model with only method, outcome, and their interaction ($\chi^2(7) = 12.50, p = .09$), nor did fit improve by only adding the main effects of proficiency and perception ($\chi^2(2) = 2.88, p = .24$). So proficiency and perception were left out of the model.

Results

Figure 2 visualizes the comparison between gains by method, task, and time. The *immediate* gains stand for changes from the pre-test to the immediate post-test, and the *delayed* gains stand for changes from the pre-test to the delayed post-test. Controlling for test task (gap or sentence), immediate gains did not differ (the average difference was only 0.05 points, 95% CI [-1.35, 1.45]) between teaching methods ($t(35) = 0.07, p = .94$), nor did delayed gains differ (the average difference was only 0.15 points, 95% CI [-0.90, 1.24]) between teaching methods ($t(35) = 0.28, p = .78$). Controlling for teaching method, immediate gains differed (on average, by 1.45 points, 95% CI [0.05, 2.85]) significantly between tasks ($t(35) = 2.09, p = .04$), but delayed gains did not differ (the average difference was only 0.85 points, 95% CI [-0.20, 1.90]) significantly between tasks ($t(35) = 1.59, p = .12$). Overall, the interaction between teaching method and task was not significant, either on the immediate post-test ($t(34) = -0.80, p = .43$), or on the delayed post-test ($t(34) = 0.47, p = .64$).

Summary of RQ2 Results

The plot in Figure 2 hints at possible interactions because the lines are not parallel, but most differences between gains are not statistically significant. The only statistically significant difference is that immediate gap task gains (the solid black line) are overall higher than immediate sentence task gains (the solid grey line), which is in concert with the answer to RQ1. Therefore, in response to RQ2, there was no difference in the efficacy of the two teaching methods. As illustrated in Figure 1, participants had a somewhat better pre-test knowledge of the items used for the hands-on condition but they showed learning and attrition rates equivalent to the hands-off condition. Finally, neither learners' proficiency nor

perceptions affected their performance gains with either DDL method.

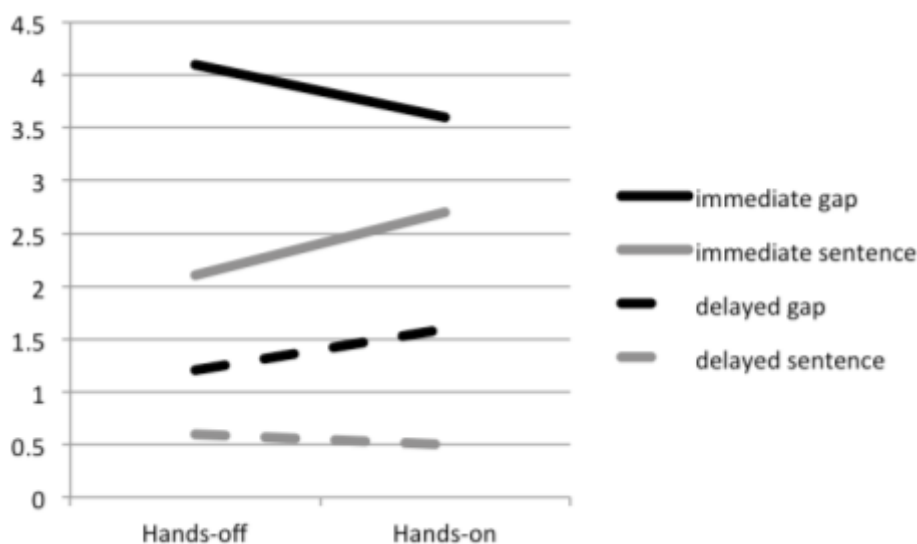


Figure 2. Average post-test gains by time, treatment method, and task.

Research Question 3

Did learners' proficiency and DDL perception change over time, and what was the relationship between proficiency, perception, and time?

Proficiency

The average L2 proficiency of the nine participants who completed both the pre-test and post-test increased over this semester-long course. More specifically, seven participants improved their individual scores and six of them moved at least half a CEFR level (e.g., from B1.1 to B1.2) or one level up. Two participants scored a few points lower on the post-course test, but their CEFR level did not change. The analysis of paired samples showed that the average group improvement of proficiency was statistically significant ($M = 9$, $SD = 10.4$, $t(7) = 2.6$, $p = .03$).

Perception

On average, learners responded positively to the questions on both pre-course and post-course perception questionnaire as evidenced by the average scores above 3 out of 5 for all questions (Table 3). Moreover, the lowest score of 1 did not appear in the responses, and seven learners out of 10 gave the highest score of 5 at least once. On average, learner DDL receptivity increased over the length of the course (Table 3, Questions 1–3), and only three learners used the low score of 2 in their post-course responses. However, the statistical analysis of paired samples showed that learner perception did not significantly change from the pre-course test to the post-course test ($M = 0.33$, $SD = 0.85$, $t(8) = 1.25$, $p = .25$). This finding can be explained by the fact that perception was already high on the pre-test. Individual differences also played a role: whereas perceptions of most (six) students increased to a different extent, it remained unchanged for two students and decreased for two students.

Proficiency versus Perception

Post-course learner preferences showed a moderately positive correlation with their post-course proficiency but this did not reach statistical significance ($r = .46$, $t(8) = 1.46$, $p = .14$).

Perception of Hands-on DDL versus Hands-off DDL

As far as the post-course reactions to the two DDL methods are concerned (Table 3, Questions 4 and 5),

the hands-off method received a slightly higher average rating, but the correlation between the hands-on and hands-off participants' receptivity was significant ($r = .68$, $t(8) = 3.0$, $p = .03$), and so was the correlation between their overall post-course receptivity and their liking of the hands-on ($r = .87$, $t(8) = 5.8$, $p = .001$) and hands-off ($r = .83$, $t(8) = 4.27$, $p = .003$) activities.

Qualitative Analysis of Learner Open-ended Commentaries

All learners indicated that they found corpus activities useful for at least some language learning purposes, including but not limited to verb-preposition collocations. The repeated theme was that corpora provide information that cannot be found in dictionaries and reference grammars. As instances of such information, students named usage frequencies and trends; stylistic norms; and examples of how words and idiomatic phrases should be used. Also revealing is what the student with the lowest L2 proficiency wrote about the difficulties experienced while working with corpus examples: “[the corpus] only really helped me if [I] knew the rest of the words, or most of them, in the examples [...] I know I made a ton of errors in the paper, but none of them would've really been fixable from [the corpus].” It is evident that he was struggling with so many other L2-related problems in the course, trying to catch up with more proficient classmates, that corpus work seemed not to have helped him in his overall achievement in the course. However, despite difficulties, this same student actually enjoyed corpus-based activities and gave fairly high ratings in the receptivity questionnaire, which may mean that the student will experiment with DDL in the future, when L2 proficiency improves. In fact, most students expressed the intent to continue using the corpus as a tool for independent learning beyond the course. Moreover, there is evidence that at least half of the participants have accomplished this: five students from this course who took another course with the researcher in the subsequent semester (that did not include compulsory DDL activities), reported having used the German corpus for independent study.

DISCUSSION

This study confirms that DDL is an effective method for teaching and learning L2 collocations (Boulton, 2010, 2012; Chan & Liou, 2005; Cobb, 1997; Daskalovska, 2015; Frankenberg-Garcia, 2014; Koosha & Jafarpour, 2006; Scripicharn, 2003; Tian, 2005; Vyatkina, 2016). In particular, it shows that DDL is effective in an area universally recognized as difficult for L2 learners across target languages: verb-preposition collocations. Indeed, the pre-test showed that intermediate L2 German learners were able to correctly supply only about 35% of the focal verb-preposition collocations despite years of instruction and an immediately preceding exposure through reading. Boulton (2010) commented on a similarly poor performance of his participants on a pre-test that targeted various L2 foci: “[t]his suggests that traditional teaching of these items is relatively unsuccessful, leaving open the possibility for alternative techniques such as DDL” (p. 543). This study's results show that this possibility was indeed successfully realized: participants both improved as a result of a DDL intervention and retained at least some of their improvement a month later. Thus, this study demonstrates that even short DDL interventions and not only prolonged and highly focused interventions (cf. Koosha & Jafarpour, 2006), which are unfeasible for most instructional settings, can be beneficial for learning verb-preposition collocations.

In regards to production tasks, participants in this study demonstrated significant immediate learning gains in both the gap-filling and sentence-writing task. However, only the gap-filling gains, not sentence-writing gains, remained significantly higher on the delayed post-test than on the pre-test. This result confirms that gap-filling tasks yield better results than freer production tasks (Cobb & Boulton, 2015) while also pointing to the need for longer or repeated interventions when targeting the development of L2 productive ability.

Hands-on and hands-off DDL were equally effective in this study, thus confirming Boulton's (2012) findings and extrapolating them to the L2 German context. This result lends support to the feasibility of using both harder and softer DDL approaches in language instruction depending on local conditions

(Boulton, 2010, 2012; Mukherjee, 2006). It appears that what made the outcomes of both approaches in this study comparable is that both DDL interventions followed the same main pedagogical principles, namely that “learning activities are student-centered and focus on language discovery” and that the instruction “relies on carefully designed and scaffolded activities” (Smart, 2014, p. 186). This approach, termed *guided induction* (or *guided discovery*), originates in general language acquisition theory (Herron & Tomasello, 1992), but has been taken up enthusiastically by DDL teachers and researchers (Flowerdew, 2009; Frankenberg-Garcia, 2014; Huang, 2008; Smart, 2014; Thomas, 2015; Yoon & Jo, 2014). In addition to teacher scaffolding, learners in this study assisted each other during the guided induction process in collaborative group work that has also been shown to be beneficial in DDL (Molés-Cases & Oster, 2015; Thomas, 2015).

The next set of results relates to the proficiency and perception effects on performance. Since Boulton (2012) is the only other study that considered all these three factors, it is worth comparing the findings. First, this study found no relationship between proficiency, perception, and DDL method (hands-on or hands-off). In contrast, Boulton (2012) found a positive correlation between proficiency and hands-off DDL performance. Furthermore, whereas Boulton found no correlation between learner preferences for hands-on and hands-off activities, this study found a strong correlation. Additionally, Boulton did not find a significant correlation between learners’ perceptions and their performance, whereas in this study, students who liked DDL improved more on the immediate gap-filling test task. Finally, learners in this study had high expectations of the new DDL method as indicated in their pre-course responses and showed no disappointment at the end of the course, as the perception outcomes did not change significantly. In contrast, the DDL receptivity of Boulton’s participants declined after the course.

Although these two studies are not directly comparable as they have some considerable design differences, some explanations may be grounded in the differences between the instructional programs. One such difference is that participants in this study were more proficient, so they had more free cognitive resources to handle the more familiar, paper-based, hands-off method and the more innovative hands-on method equally well and they also liked them equally much (or equally not very much). Another important difference is the motivation factor: whereas Boulton’s (2012) participants took their language course as a requirement, all students in this study were highly motivated and, therefore, they appreciated a new method to improve their L2 knowledge. The power of motivation and importance of learner involvement has been pointed out both in DDL research (e.g., Daskalovska, 2015) and general SLA research (e.g., Hulstijn & Laufer, 2001). In this respect, the students in this study appear to be similar to those in a study by Smart (2014), who states, “The DDL learners’ engagement with the material and interest in what many perceived as a novel approach to grammar instruction may have led to more learning during the course of this brief instructional intervention” (p. 196).

When compared with a study that included a delayed post-test, this study corroborates the findings by Chan & Liou (2005) that lower-proficiency learners benefit more from DDL in the short term but this advantage disappears in the long term. Additionally, this advantage was found in this study only on the gap-filling task, whereas more proficient learners improved more than less proficient learners on the sentence-writing task (which was not employed by Chan and Liou, so no comparison is possible). This result shows that combining a new teaching method (DDL) and a cognitively challenging task may be more appropriate for more proficient learners (in the upper score range at B1 CEFR level and above B1). This result can be also explained by the ceiling effect as more proficient learners already had relatively high scores on the pre-test gap-filling task, whereas they had more room for improvement on the sentence-writing task. In the long term, proficiency and perception made no difference for learner performance on either task. In this study, the sentence-writing task was especially difficult as it involved supplying accurate noun morphology. The fact that no residual effect was found on the sentence-writing delayed post-test in comparison with the pre-test points to the necessity of recycling this complex L2 German feature at later times in the instructional sequence.

CONCLUSION AND IMPLICATIONS FOR FUTURE RESEARCH

In sum, the results of this study are highly encouraging and show that different DDL types (computer-based and paper-based) are equally effective for teaching L2 collocations and benefit learners at different proficiency levels (see also Vyatkina, 2016), provided that they follow a guided induction approach with sufficient instructor and peer scaffolding. While showing that even one-time DDL interventions are beneficial for short-term learning and for production on highly controlled tasks (gap-filling), the results also point to the necessity of prolonged or repeated interventions for achieving long-lasting gains in freer L2 production and in learning complex linguistic features such as inflectional morphology. Last but not least, this study, along with Vyatkina (2016), expands the DDL target language landscape to L2 German.

These results still need to be interpreted with caution, however, as the number of participants was small. More research is needed to fully account for the effects of different DDL activities as applied to various learning contexts and target structures. In particular, this study's participants were fairly motivated L2 learners at intermediate proficiency levels with a DDL researcher acting as their instructor. Future studies should explore what types of DDL interventions administered by real-life instructors who are not corpus researchers work with learners with lower proficiency and motivation. With regard to research methodology, DDL scholars should consider multifactorial designs in order to systematically explore the influence of multiple factors on L2 learning.

Finally, this study showed broader positive effects of integrating DDL and non-DDL components in a language course as learners' overall proficiency increased significantly. As Kennedy and Miceli (2010, p. 29) have argued, DDL should be implemented as "apprenticeships" in "mastering corpus consultation as a gradual, long-term process" to have broader and long-lasting effects. This study was designed as such an apprenticeship that supported development of DDL skills in learners. Only future longitudinal studies can explore whether and how the broader DDL potential can be realized. However, the following learner comments indicate that they appreciated DDL and that it may indeed become a life-long L2 learning tool for them:

- "I think this skill set is the most important thing I've learned in the course."
- "[The corpus] is one of the most useful resources we used for learning German."
- "I believe the corpus was useful in solving the many mysteries of German grammar."
- "I found the [corpus] resource enormously useful as a student, trying to gain a more nuanced understanding of German's linguistic patterns, in a way that dictionaries and even grammar books don't often show."

APPENDIX A. List of Target Verb-preposition Collocations with English Translations

Collocations targeted in the hands-on DDL module:

<i>sich beschäftigen mit</i> + dat	to occupy oneself with, to engage in
<i>duften nach</i> + dat	to smell of
<i>erinnern an</i> + acc	to remind of
<i>sich gewöhnen an</i> + acc	to accustom oneself to
<i>sich handeln um</i> + acc	to be about
<i>sich kümmern um</i> + acc	to concern oneself with
<i>sorgen für</i> + acc	to care for
<i>träumen von</i> + dat	to dream about
<i>verzichten auf</i> + acc	to do without, to relinquish
<i>warten auf</i> + acc	to wait for

Collocations targeted in the hands-off DDL module:

<i>sich bemühen um</i> + acc	to go after
<i>hoffen auf</i> + acc	to hope for
<i>schmecken nach</i> + dat	to taste of
<i>greifen nach</i> + dat	to catch at, grab for
<i>verfügen über</i> + acc	to have at one's disposal
<i>Wert legen auf</i> + acc	to attach importance to
<i>Es geht um</i> + acc	to be about [it is about]
<i>deuten auf</i> + acc	to point at / to
<i>bestehen aus</i> + dat	to consist of
<i>suchen nach</i> + dat	to search for

APPENDIX B. Example of a Test Worksheet (English translations added to directions)

A. Füllen Sie die Lücken in den folgenden Sätzen aus *Blaubär* mit den richtigen Präpositionen. (Fill the gaps in the following sentences from *Bluebear* with correct prepositions.)

Z.B. (e.g.): *Plötzlich mußte ich an die Klautergeister **denken**.*

Wenn man relativ unerfahren auf See ist und den Horizont beobachtet, dann glaubt man, jeden Augenblick müsse irgend etwas Atemberaubendes an ihm zum Vorschein kommen. Aber das einzige, was dahinter _____ einen **wartet**, ist ein neuer Horizont.

Qwert hingegen liebte das Geräusch, weil es ihn _____ einen populären Schlager aus der 2364. Dimension **erinnerte**.

Ich **beschäftigte mich** _____ den letzten Fragen der Philosophie, aber diese simple Frage hatte ich mir noch nicht gestellt.

Mac **sorgte** _____ die passende Ernährung.

Seien Sie lieber demnächst etwas vorsichtiger und **verzichten** Sie in Zukunft _____ fleischliche Ernährung!

Es **handelte sich** _____ einen der ortstypischen Zuchthöfe, seine Besitzer waren vermutlich Hals über Kopf geflohen und hatten ihre Schützlinge hilflos und gefangen der Gefahr überlassen.

Während Mac in der Gegend herumflog, um andere Rettungssaurier zusammenzutrommeln, die uns helfen sollten, die Welpen in ein anderes Heim zu bringen, **kümmerte ich mich** _____ die drolligen Tierchen.

Im Wald wucherten neuerdings unterirdisch dicke, heftig duftende Trüffelpilze. Ich mußte mich _____ ihren intensiven Geschmack **gewöhnen**, aber sobald das geschehen war, konnte ich nicht mehr von ihnen lassen.

Ich dämmerte nur noch ein wenig und **träumte** im Halbschlaf _____ der nächsten Mahlzeit.

Dünne Lianenfäden hingen von Bäumen herunter, **dufteten** sanft _____ Knoblauch und ließen sich schlürfen wie Spaghetti.

B. Schreiben Sie je einen kurzen Satz mit jedem angegebenen Verb und einer passenden Präposition. (Write short sentences with each verb and a corresponding preposition.)

1. warten:

2. sich beschäftigen:

3. sorgen:
4. sich gewöhnen:
5. träumen:

APPENDIX C. Hands-on DDL Worksheet (1st Page): Instructions, Model, and the 1st Task (English translations added)

Einige deutsche Verben werden mit einer bestimmten Präposition und mit einem bestimmten Kasus gebraucht. Recherchieren Sie einige Verben in DWDS. 1) Kopieren Sie 2 Beispielsätze mit (Verb + Präposition), 2) unterstreichen Sie die Präpositionalphrase (PP) und 3) schreiben Sie, in welchem Kasus diese Phrase steht.

(Some German verbs are used with a certain preposition and a certain case. Research some verbs in DWDS. 1) Copy 2 example sentences with (verb + preposition), 2) underline the prepositional phrase (PP), and 3) write the case of this phrase.)

Beispielsanfrage (sample query): *sehnen sich (to crave)*

1. Xerxes hat seine Geliebte Amastris verlassen und sehnt sich nach einer neuen Liebe.
2. Wir sehnen uns nach diesem Frieden.

Antwort (answer): *sich sehnen nach + Dativ*

Aufgaben (tasks):

1. Anfrage (query): *beschäftigen sich (to occupy oneself with)*

Beispiele (examples):

- 1.
- 2.

Antwort (answer):

sich beschäftigen _____ + _____

APPENDIX D. Hands-off DDL Worksheet (1st Page): Instructions, Model, and the 1st Task (English translations added)

Unterstreichen Sie die Präpositionalphrase (PP), die nach dem Verb in den DWDS-Konkordanzen steht, und schreiben Sie, in welchem Kasus diese Phrase steht.

(Underline the prepositional phrase (PP) that follows the verb in the DWDS concordances and write the case of this phrase.)

Beispiel: *sich sehnen*

(Example: *to crave*)

Bedeutung: innig, schmerzlich nach jmdm., etw. verlangen, jmdn., etw. herbeiwünschen

(Meaning: to long, yearn for so. / sth.)

DWDS Kernkorpus-Konkordanzen (DWDS core corpus concordances):

1	[1999]	[B]	duecker	Ich sehnte mich nach Gregor , und ich sehnte mich auch noch nicht wieder
2	[1999]	[G]	schwani	Sie sehnten sich nach etwas Lebendigem , Heiterem und Natürlichem .
3	[1999]	[Z]	zeit199	Und » fast alle « sehnten sich nach den leeren Regalen , sie wollten fort in die gute alte
4	[1999]	[Z]	zeit199	Sie sehnen sich nach dem Lafontaine zurück , den sie brauchen .
5	[1998]	[G]	kellner	; und das , was von der guten alten Post noch übrig ist , sie alle sehnen sich nach » Charismatikern « . Leider handelt es sich dabei nicht
6	[1998]	[W]	oper	Nur Galatea stimmt nicht in den Jubel ein , denn sie sehnt sich nach Acis .
7	[1998]	[W]	oper	in.) . 1. Akt. Xerxes hat seine Geliebte Amastris verlassen und sehnt sich nach einer neuen Liebe (Ombra mai fu) . Romilda erregt sie
8	[1998]	[Z]	zeit199	heime und Ostelend ; der nicht mehr junge Wilde Frank Castorf sehnt sich nach Menschen , die etwas für ihre Überzeugungen riskieren

Antwort (answer): *sich sehnen nach* + Dativ

Aufgaben (tasks):

1. *sich bemühen* (to go [after])

Bedeutung: sich anstrengen; sich um jmdn. kümmern

(Meaning: to exert oneself, try for sth.)

DWDS Kernkorpus-Konkordanzen (DWDS core corpus concordances):

9	[1983]	[Z]	adg1983	Wir bemühen uns um einen Friedenszustand für ganz Europa , der es uns
10	[1981]	[W]	haberma	In anderen Fällen ist der Erwerb optativ - nur diejenigen bemühen sich um ein eigenes Totem , die sich aus dem Kollektiv heraus
11	[1979]	[W]	mgg1979	er Volloper usw. Verschiedene Spezialarbeiten aus jüngster Zeit bemühen sich um die Gewinnung eines fester umrissenen Bildes .
12	[1979]	[Z]	adg1979	ich Durchsetzung und Verwirklichung der Menschenrechte ; wir bemühen uns um einen konstruktiven Beitrag unserer Länder und der
13	[1975]	[Z]	adg1975	Einige der hier vertretenen Regierungen bemühen sich um eine Einigung auf einem noch wichtigeren Gebiet mit
14	[1974]	[Z]	adg1974	Andere europäische Länder bemühen sich um den Wiederaufbau einer freiheitlich-demokratischen
15	[1974]	[Z]	adg1974	, der folgenden Wortlaut hat : " Art. 1 : Die beiden Regierungen bemühen sich um die Festigung und Entwicklung der Bande der auf

Antwort (answer):

sich bemühen _____ + _____

NOTES

- In addition to simple verbs (*warten* [to wait]), the list included some phrases (*Wert legen* [to attach importance]) and reflexive verbs (*sich beschäftigen* [to occupy oneself]).

ACKNOWLEDGEMENTS

This study was supported in part by the Department of Germanic Languages and Literatures of the University of Kansas. I would also like to acknowledge Terrence D. Jorgensen Jr. for his invaluable help with the statistical analysis.

ABOUT THE AUTHOR

Nina Vyatkina is Associate Professor of German and Applied Linguistics and Chair of the Department of Germanic Languages and Literatures at the University of Kansas. Her research interests include instructed Second Language Acquisition, data-driven learning, learner corpus analysis, and longitudinal learner language development.

E-mail: vyatkina@ku.edu

REFERENCES

- Barfield, A. (2013). Lexical collocations. In C. A. Chappelle (Ed.), *The encyclopedia of applied linguistics*. Oxford, UK: Wiley-Blackwell.
- Baten, K. (2011). Processability theory and German case acquisition. *Language Learning*, 61(2), 455–505.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165–182). Amsterdam, Netherlands: Rodopi.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, UK: Pearson Education.
- Boulton, A. (2010). Data-driven learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534–572.
- Boulton, A. (2012). Hands-on / hands-off: Alternative approaches to data-driven learning. In J. Thomas & A. Boulton (Eds.), *Input, process, and product: Developments in teaching and language corpora* (pp. 152–168). Brno, Czech Republic: Masaryk University Press.
- Boulton, A., & Pérez-Paredes, P. (Eds.). (2014). Researching new uses of corpora for language teaching and learning [Special issue]. *ReCALL*, 26(2).
- Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language*, 82(4), 711–733.
- Carter, R., & McCarthy, M. (1995). Grammar and the spoken language. *Applied Linguistics*, 16(2), 141–158.
- Chambers, A. (2005). Integrating Corpus Consultation in Language Studies. *Language Learning & Technology*, 9(2), 111–125. Retrieved from <http://llt.msu.edu/vol9num2/chambers/>
- Chan, P.-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.
- Cobb, T. (1997). Is there any measurable learning from hands-on concordancing? *System*, 25(3), 301–315.
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics*. Cambridge, UK: Cambridge University Press.
- Council of Europe. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Strasbourg, France: Language Policy Unit. Retrieved from http://www.coe.int/t/dg4/linguistic/source/framework_en.pdf
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369–382.
- Daskalovska, N. (2015). Corpus-based versus traditional learning of collocations. *Computer Assisted Language Learning*, 28(2), 130–144.
- De Felice, R., & Pulman, S. (2009). Automatic detection of preposition errors in learner writing. *CALICO Journal*, 26(3), 512–528.

- Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290–325.
- Ellis, N. C. (2014). Cognitive AND social language usage. *Studies in Second Language Acquisition*, 36(3), 397–402.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393–417.
- Flowerdew, L. (2015). Data-driven learning and language learning theories: Whither the twain shall meet. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 15–36). Amsterdam, Netherlands: John Benjamins.
- Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, 26(2), 128–146.
- Frankenberg-Garcia, A., Flowerdew, L., & Aston, G. (Eds.). (2011). *New trends in corpora and language learning*. New York, NY: Continuum.
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells? *TESL-EJ*, 8(4), 1–37.
- Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam, Netherlands: John Benjamins.
- Gavioli, L. (2009). Corpus analysis and the achievement of learner autonomy in interaction. In L. Lombardo (Ed.), *Using corpora to learn about language and discourse* (pp. 39–71). Bern, Switzerland: Peter Lang.
- Herron, C., & Tomasello, M. (1992). Acquiring grammatical structures by guided induction. *The French Review*, 65(5), 708–718.
- Huang, L.-S. (2008). Using guided, corpus-aided discovery to generate active learning. *English Teaching Forum*, 46(4), 20–27.
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558.
- Johns, T. (1990). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14–34.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning & Technology*, 5(3), 77–90. Retrieved from <http://lt.msu.edu/vol5num3/kennedy/>
- Kennedy, C., & Miceli, T. (2010). Corpus-assisted creative writing: Introducing intermediate Italian students to a corpus as a reference resource. *Language Learning & Technology*, 14(1), 28–44. Retrieved from <http://lt.msu.edu/vol14num1/kennedymiceli.pdf>
- Koosha, M., & Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: The case of Iranian EFL adult learners. *Asian EFL Journal Quarterly*, 8(4), 192–209.
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 11–23). Harlow, UK: Longman.
- Leńko-Szymańska, A., & Boulton, A. (2015). *Multiple affordances of language corpora for data-driven learning*. Amsterdam, Netherlands: John Benjamins.
- McEnery, T., Xiao, R., & Tono, Y. (Eds.). (2006). *Corpus-based language studies: An advanced resource book*. London, UK: Routledge.

- Moers, W. (2002). *Die 13^{1/2} Leben des Käpt'n Blaubär*. Munich, Germany: Goldmann.
- Molés-Cases, T., & Oster, U. (2015). Webquests in translator training: Introducing corpus-based tasks. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 199–224). Amsterdam, Netherlands: John Benjamins.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, K. Kohn, & J. Mukherjee (Eds.), *Corpus technology and language pedagogy: New resources, new tools, new methods* (pp. 5–24). Frankfurt am Main, Germany: Peter Lang.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223–242.
- Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130–149.
- Robinson, P. (1995). Attention, memory, and the “noticing” hypothesis. *Language Learning*, 45(2), 283–331.
- Robinson, P., & Ellis, N. C. (Eds.). (2008). *Handbook of cognitive linguistics and second language acquisition*. London, UK: Routledge.
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205–225.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge, UK: Cambridge University Press.
- Scripicharn, P. (2003). Evaluating classroom concordancing: The use of corpus-based materials by a group of Thai students. *Thammasat Review*, 8(1), 203–236.
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15(2), 165–179.
- Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26(2), 184–201.
- Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, 63(1), 121–159.
- Sun, Y.-C., & Wang, L.-Y. 2003. Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, 16(1), 83–94.
- Szudarski, P., & Carter, R. (2016). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics*, 26(2), 245–265.
- Thomas, J. (2015). Stealing a march on collocation: Deriving extended collocations from full text for student analysis and synthesis. In A. Leńko-Szymańska & A. Boulton (Eds.), *Multiple affordances of language corpora for data-driven learning* (pp. 85–108). Amsterdam, Netherlands: John Benjamins.
- Thomas, J., & Boulton, A. (Eds.). (2012). *Input, process, and product: Developments in teaching and language corpora*. Brno, Czech Republic: Masaryk University Press.
- Tian, S. (2005). The impact of learning tasks and learner proficiency on the effectiveness of data-driven learning. *Journal of Pan-Pacific Association of Applied Linguistics*, 9(2), 263–275.

- Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(2), 207–226.
- Yoon, H., & Jo, J. (2014). Direct and indirect access to corpora: An exploratory case study comparing students' error correction and learning strategy use in L2 writing. *Language Learning & Technology*, 18(1), 96–117. Retrieved from <http://lt.msu.edu/issues/february2014/yoonyjo.pdf>