COLLABORATION TWO-WAY: WORKLOAD AND CO-OWNERSHIP IN L2 WIKI WRITING

Nike Arnold
Portland State University

Lara Ducate
University of South Carolina

Claudia Kost
University of Alberta

ABSTRACT

Collaborative writing has been found to lead to more productive writing processes and enhanced final products in terms of a richer vocabulary, more accurate grammar, and better organization. The present study expands on this research strand by exploring if different group writing processes affect the quality of wiki texts composed by groups of intermediate German L2 learners. Defining true collaborative writing as involving both a balanced workload and a joint responsibility for the product from all group members, it measured collaboration in two ways. Results indicate that most of the 19 groups in this study had a somewhat unbalanced workload with wide variability in editing group members' contributions. Although the wiki texts differed greatly with regard to quantitative measures of length, accuracy and cohesion, no correlation was found in terms of workload or co-ownership. While holistic ratings of the texts concerning accuracy and cohesion seemed at times incongruent with the analytic measures, the raters' comments provided a perspective that captured facets and nuances of a text that the analytic indicators did not.

INTRODUCTION

Rooted in socioconstructivist pedagogies, online and paper collaborative writing has been researched at various proficiency levels with multiple task types and has been shown to provide advantages in terms of the writing process and the final product. Studies have found that collaborative writing can lead to a richer vocabulary (Amir, Ismail & Hussin, 2011; Shehadeh, 2012), more accurate grammar (Amir, Ismail & Hussin, 2011; Storch, 1999, 2005; Wigglesworth & Storch, 2009), higher level writing (Mak & Coniam, 2008; Storch, 2005), and better organization (Aydin & Yildiz, 2014; Shehadeh, 2012). When learners work together on texts, they often take more ownership when providing feedback and enjoy receiving quicker feedback, which possibly

leads to better products than if they had worked alone or had just provided peer feedback (Storch, 2005). Peer editing and collaborative writing also allow writers to pool ideas and resources and encourages them to work together towards a better product (Aydin & Yildiz, 2014). This collaboration can also promote critical thinking as students feel encouraged to organize their thoughts and ideas to provide useful feedback for the success of the group's final product (Aydin & Yildiz, 2014; Oskoz & Elola, 2012; Limbu & Markauskaite, 2015). Furthermore, students have reported that collaboration helped them to enjoy the assignment more and assisted in their development of a sense of community (Elola & Oskoz, 2010; Lin & Maarof, 2013; Lund, 2008; Storch, 2005; Xuanxi, Chu, & Ki, 2014).

Reviewing these various benefits of collaborative writing, however, one might wonder about the exact nature of collaboration. When learners "work together" on a task, groups might complete it in very different ways. This ultimately raises an important question: Do different collaborative processes affect the quality of the resulting product? This study examines intermediate learners' wiki-mediated collaboration from two different angles, workload and co-ownership, in order to assess how the process influences the product.

Collaboration vs. Cooperation

While not all studies on collaborative writing have distinguished between collaboration and cooperation (Amir, Ismail & Hussin, 2011) or even established how learners worked on a group task, the current study sees a clear demarcation between the two terms. Collaboration refers to a task taken on by two or more people that is completed together, where all participants are involved in and take responsibility for every aspect of the task, and everyone shares decision-making regarding the final product (Allen, Atkinson, Morgan, Moore, & Snow, 1987; Alyousef & Picard, 2011; Dillenbourg, 1999; Yang, 2014). The collaborative process is a social one that requires negotiation, coordination, and communication between all group members throughout the process (Elola & Oskoz, 2010; Kessler & Bikowski, 2010; Lin & Maarof, 2013; Lowry, Curtis, & Lowry, 2004; Storch, 2005; Yeh, 2014). Cooperation, on the other hand, implies that group members complete separate portions of the project, work independently from the rest of the group and combine their work at the end; the same degree of negotiation and communication is not necessary in cooperation. It is also possible that groups participate in a mix of collaboration and cooperation. Based on these definitions, this study proposes that true collaborative writing involves the following: 1) an equal workload among group members, and 2) all members taking joint responsibility for the text as a whole. These two indicators form the basis for how collaboration was measured for the present study to establish any relationship between process and product.

Although instructors sometimes have clear ideas about how they want their students to approach collaborative and cooperative writing tasks, learners do not necessarily view it the same way. Limbu and Markauskaite (2015) found that students perceived what they termed collaborative writing in four different ways. Some students viewed it as a division of work for the sake of efficiency and others saw it as a way to produce a better final product by combining different levels of expertise. A third group engaged in collaboration to reach a deeper understanding of the course and writing content, while a fourth group focused more on the affective side of collaboration and used it to develop better skills for cooperation and interaction. Depending on how they viewed

collaboration, learners also had different ways of interpreting who controlled the collaborative space. Some saw it as a space organized by the teacher, others viewed it more as a space guided by the members and their interactions, and the final group saw it even more openly as a space created and maintained by the learners.

While students sometimes have different understandings of collaboration, it has also been found that students prefer cooperative learning to collaborative writing (Alyousef & Picard, 2011), possibly because they feel it is easier to work alone than with others, even though the final product might not be as good. These different ways in which students perceive group work illustrate that there is no consensus among students regarding what constitutes collaboration. Since the definitions of collaboration in the literature as well as for students are somewhat cloudy, for the purpose of this study, the term 'group writing' will be used as an umbrella term for any type of writing that involves two or more students.

Group Writing

As mentioned above, group writing can take place on paper or online. Computer-mediated group writing can easily be facilitated with tools such as a wikis (Arnold, Ducate, & Kost, 2012; Aydin & Yildiz, 2014; Elola & Oskoz, 2010; Kessler, 2009; Kessler & Bikowski, 2010) and online word processing (Kessler, Bikowski, & Boggs, 2012). Given technology's unique affordances for collaboration and writing, such writing might be fundamentally different from paper-based tasks, as supported by the findings of Martinsen and Miller (2012). As one of the few studies comparing group writing on wikis vs. paper, they found that wikis encouraged more true collaboration among beginning foreign language learners, whereas paper-based tasks encouraged more cooperation with students rarely reading each other's work before assembling it. Learners did, however, appreciate the face-to-face communication that was available for the paper composition, which led the authors to conclude that collaboration, where all learners contribute to each part of the composition, should be encouraged regardless of the tools used.

Levels of collaboration have been shown to be affected by a variety of factors, such as L1 or L2 proficiency or task type. Yang's study (2014) examined group writing with ESL business students at two Canadian universities. Three groups of students completed several group writing assignments, but it is unclear from the study's description how exactly the texts were written (online, computer-based, or on paper). Findings indicate that the success of the groups' collaboration was largely based on their L1 background, L2 proficiency, and the rules established within their groups with regard to task division, working together, getting along, and taking control. Members of a group with the same L1 could communicate and mediate with each other more effectively, while members with more advanced L2 proficiency were found to split the work more evenly than those with lower L2 proficiency. Students' backgrounds also played a role in the effectiveness of their collaboration. If they were more used to working with a group, they were more effective and productive.

The effect of task type was illustrated in Aydin and Yildiz's (2014) study of EFL university students in Turkey who participated in three wiki-based group writing tasks: the argumentative task encouraged more peer-corrections than the informative and decision-making tasks, while the

informative task received the most self-corrections. The authors suggested that the argumentative and decision-making tasks were more difficult and required more collaboration than the informative task, which students felt more comfortable completing alone. The grammatical structures were 94% correct when the corrections were considered for all of the tasks and the students focused more on meaning than form in every task, similar to the findings of other studies (Arnold, Ducate, & Kost, 2009; Kessler, 2009; Kessler, Bikowski, & Boggs, 2012).

While proficiency, task type, and previous experience with group writing have been found to influence the degree of collaboration, another strand of research has investigated the benefits of group writing. Focusing on traditional paper-based writing, Storch's (2005) analysis of students' dialogues during a paper-based writing task illustrated that learners discussed various aspects of the writing process and were able to help each other generate ideas, learn about different perspectives, and provide each other with feedback (Storch, 2005). Elola and Oskoz (2010) focused on asynchronous wiki writing tasks complemented with synchronous group chats, and also concluded that the collaborative groups in their study were more successful in their writing than the individuals due to their planning and revision practices. The transcripts from the group chats, for example, indicated that students were engaged in discussion about their essays, which could have ultimately improved the writing process and given them more strategies for writing their essays. There were, however, no statistical differences in fluency, accuracy, or complexity between group and individual writings. Yeh (2014) found analogous results when examining the essays produced by groups with higher proportions of collaborative dialogue, median proportions, and lower proportions in a synchronous online writing environment. Similar to Storch (2005) and Elola and Oskoz (2010), the groups that were more collaborative in their communication produced better essays in terms of fluency and accuracy, although all groups benefited from the idea generation, discussion about the writing process, and editing discussion. From these studies, it seems that collaboration throughout the writing process benefits both the generation and organization of ideas as well as the final product.

Strobl's (2014) findings, however, do not support this conclusion. Comparing individually and jointly written online texts from advanced L2 learners, she reported no difference in terms of complexity, accuracy, coherence or cohesion, although the collaborative texts were more effective in regard to content selection and organization due to the groups' detailed discussions during planning. Unlike Storch's (2005) findings for paper writing, the collaborative texts were longer, possibly due to the fact that several authors' texts were combined and that they were written online. The difference in accuracy was possibly due to the fact that all of the participants in the Strobl (2014) study had an advanced proficiency level and therefore did not engage in as much scaffolding regarding accuracy or that students wrote online rather than on paper. Unfortunately, there were some free-riders in Strobl's study, which contributed to a complete break-down in one of the groups and one of the group members ultimately decided to complete the assignment individually. The free-ride phenomenon has also been documented by Arnold, Ducate, and Kost (2012), who discovered that some students did not equally participate in the writing process.

In addition to examining the benefits of group writing, other studies have researched students' attention to accuracy and meaning during the writing process. Kessler, Bikowski, and Boggs (2012) observed ESL Fulbright scholars' online writing practices and products when completing a joint

research project and found that they focused more on meaning than form, which differs from other studies in which learners focused on form over meaning and were able to correct most of their errors together (Arnold, Ducate, & Kost, 2009; Strobl, 2014). Similar to the findings of Arnold, Ducate, and Kost (2012) and Strobl (2014), they also did not all participate equally. The researchers posited that some students were perhaps unfamiliar with the web-based word processing program or group writing in general, did not appreciate the roles assumed by their team members, or were uncomfortable writing in English. The students, however, did not mention an issue with the varying rates of participation. In another wiki group writing study, Kessler (2009) also found that EFL teacher candidates focused more on meaning than form and overlooked many errors that did not interfere with meaning. Participants felt comfortable editing their peers' posts and were confident in the group task. In a follow-up analysis of the same project, Kessler and Bikowski (2010) investigated the group behavior further and found that not everyone participated equally. The work seemed to fall into three phases, including "build and destroy, full collaboration, and informal reflection" (p. 48). Only a few students participated in the first phase when they were quickly adding and deleting content. More students collaborated during the second phase, and during the third phase students followed another student's lead and included their own personal reflections on the course. At the end, 18 students who had not yet participated made cosmetic changes, suggesting that they were perhaps either just reading throughout the process or that they simply wanted to fulfill the requirement to participate at the end. As illustrated by the studies above, a risk in group writing is that not all students might participate equally, which can affect group dynamics and students' overall impressions of working together. Furthermore, such uneven workload might affect the quality of the product – an issue addressed in the present study.

The purpose of the following study was to investigate the connection between the process of L2 collaborative wiki writing and its final product. More specifically, it addresses the following research questions:

- When small groups of intermediate foreign language learners write a wiki together as a class assignment, how many groups work collaboratively?
- Is the length of the wiki texts related to the degree of collaboration within the groups?
- Is the quality of the wiki texts in terms of accuracy and cohesion related to the degree of collaboration within the groups?

To reflect how collaboration has been defined in the literature, two different operationalizations were viewed as central indicators of true collaboration for this study: balanced workload distribution among group members and members taking co-ownership of the text as a whole. Given the shortage of research on L2 collaborative wiki writing, the present study is exploratory in nature and was not based on any a priori hypotheses.

METHODOLOGY

Participants & Context

The participants were 53 learners of German as a foreign language enrolled in three intact intermediate German classes at three public universities in North America (Class 1: n=25; Class 2: n=10; Class 3: n=18), which were taught by the three researchers. All participants were native speakers of English in their late teens or early twenties.

The German courses were based on communicative curricula focused on interpretive, interpersonal and presentational communication and included other formal writing assignments in addition to the wiki. The Class 1 was a 4th semester course, in which most students had enrolled to fulfill the university's foreign language requirement. Classes 2 and 3 were 5th semester composition & conversation courses for majors or minors. Despite the different levels of the courses, all of the students had received approximately the same number of instructional hours before the study.

Intermediate level learners were selected for this study because learners at this level are able to "create with the language and communicate simple facts and ideas in a series of loosely connected sentences" (ACTFL, 2012, p. 13) and are therefore beginning to produce longer and more complex written discourse, an important skill for this type of assignment. In addition, intermediate learners have varying linguistic strengths, such as a stronger vocabulary or better understanding of grammar, which should make collaboration particularly valuable.

Task

As one of the main writing assignments, the wiki task was directly linked to the course reading, a graded reader of the novel *Am kürzeren Ende der Sonnenallee* by Thomas Brussig (2003). Set in socialist East Berlin in the 1970s, the plot includes frequent cultural and historical references with which North American students of that generation are unlikely to be familiar. Since this lack of background knowledge could affect reading comprehension, the wiki was intended to function as a resource about topics such as the Stasi secret police, censorship and the Socialist party (see tables in the results section for a complete list of topics). Students worked outside of class in self-selected groups of 2-4 to research their assigned topic and compose a wiki page summarizing the most relevant information. There was a total of 19 groups: five pairs, 13 triads and one group of four.

While the overall goal of the wiki assignment was the same in all three classes, its implementation differed somewhat, namely with regard to timing and structuring of the task. Classes 2 and 3 completed the wiki as a pre-reading task with separate deadlines for an annotated bibliography, outline, and first and final drafts with received feedback from the instructor and/or peers throughout these steps. The implementation in Class 1 occurred after students had read the novel (extension task) and was less structured: students received instructor feedback upon completion of the wiki, which was also the basis for an oral presentation. Despite these differences, however, the overall purpose and design of the assignment was similar.

A wiki was chosen for this writing assignment for the following reasons: 1) its space- and time-

independent nature facilitates out-of-class collaboration, 2) it becomes a published document that everyone can easily access, which provides an audience and matches learners' established practices for online information gathering, 3) its archives provide access to old versions, which can encourage editing and revising, and 4) its archived versions provide a level of accountability for the students. The wiki pages of Class 1 were hosted on PBWiki while Classes 2 and 3 worked with Wikispaces.

Analysis

Independent Variable: Collaboration

As mentioned above, the independent variable of collaboration was operationalized in two ways: workload and co-ownership. Unable to rely on previous research, we devised ways to numerically capture both operationalizations for this study. The wiki was used to compare all archived page versions and identify the edits, their authors and the author of the original text. In the example in Figure 1, a group member who did not write the original text added all of the bolded words or phrases after deleting the crossed-out sections. This student found what s/he thought were mistakes and then changed them to what s/he considered more accurate grammar.

Figure 1.
Wiki output identifying deletions and additions between two text versions.

Die Grenze

Nach dem Krieg wurde Deutschland abgeteilt. Deutschland ~~wird~~ **wurde** in 4 ~~Zone~~ **Zonen** abgeteilt. Berlin wurde auch in 4 ~~Zone~~ **Zonen** abgeteilt. Die Leute von Ost Berlin ~~wollte~~ **wollten** nach West Berlin kommen. Die DDR Regierung wollte, dass die Menschen in Ost Berlin bleiben. Die Berliner Mauer hat die Freiheit zwischen die ~~Ost-Berliener und West-Beliener abteilen~~ **Ost-Berliner und West-Berliner abgeteilt**. Alle ~~wollte Berlin kontrolliert~~ **wollten Berlin kontrollieren**.

Hintergrund

Die Berliner Mauer war eine Grenze zwischen Ost und West Berlin. Ab dem 23.8. 1961 durften West-Berlin ~~Bürgeren~~ **Bürger** Ost-Berlin nicht mehr betreten. Die DDR wollte, dass die Menschen in Ost-Berlin bleiben. Am August 13. 1961 hat die Bauwerke von die dem Grenze begonnen. Die Berliner Mauer hat die Berliner Leute eingefangen. ~~10,000~~ **10.000** Ost-Berliner Leute ~~hat probiert~~ **hatten versucht** zu Entkommen. Aber nur ~~5,000~~ **5.000** Leute ~~hat~~ **hatten** es gemacht. Am 17.8.1962 wurde der achtzehnjährige Ost-Berliner Peter Fechter bei einem Fluchtversuch über die Mauer von Ost-Berliner Grenzwachen angeschossen, und er verblutete.

To numerically capture the **workload** distribution within a group, the following process was used:

1. Each group member was assigned a numerical value based on the percentage of edits made to all archived versions of the page (see Table 1 below). The highest number was assigned to students who were categorized as team players and did their fair share within the group. In a group of three, for example, a balanced workload meant that every member did about a third of the work. If a member did significantly more or less work, the workload

distribution was uneven. Therefore, such roles received a lower numerical value.

2. The numerical values for all group members were added up and divided by the number of members to account for different groups sizes.

Table 1
Coding of Group Member Roles

Numerical Value	**Percentage of Edits Performed by Individual Member**		
	Groups of 2	Groups of 3	Groups of 4
0	< 10%	< 10%	< 10%
1	< 40%	<25%	<20%
2	40-60%	25-35%	20-30%
1	> 60%	> 40%	> 30%

As a result, a group's workload index can range from 0.25-2.0. The highest possible index, 2.0, represents a group with a balanced workload distribution, meaning that all members contributed their fair share.

While the workload index was calculated based on members' total edits, collaboration as **co-ownership** examines the location of edits: Did members edit their own contributions or those of their peers? The co-ownership index was calculated as follows:

1. Based on the wiki page archives, the total number of edits members made in sections of the text that someone else had originally added was calculated.
2. This number was then divided by the total number of edits the group performed on the text.

With a possible range of 0% to 100%, this number represents the percentage of edits the group made in sections originally written by another group member. In other words, a higher co-ownership index indicates that members took more responsibility for the text as a whole. A low number, in contrast, indicates that members focused their edits mainly on their own section and engaged in cooperation by dividing the work into sub-tasks to be assembled into a larger whole (Dillenbourg, Baker, Blaye, & O'Malley, 1996; Haythornthwaithe, 2006).

Dependent Variables: Length, Accuracy and Cohesion

The final versions of groups' wiki pages were analyzed based on three dimensions: length, accuracy and cohesion. There were multiple rationales for choosing these dimensions. As illustrated in the literature review, these measures have traditionally been used in L2 writing research and can be more easily quantified for statistical analyses than other measures (e.g., content). Furthermore, these dimensions seem particularly relevant to the time- and space-independent collaborative writing that wikis encourage, particularly length and cohesion. Lastly, analyzing for length, accuracy and cohesion provides a multidimensional view of the texts that these intermediate L2 learners produced.

Length was measured by the total number of words that comprised the final version of the wiki page. That included titles, section headings and captions for visuals. Excluded were lists of

references, vocabulary lists, long quotes from other sources (e.g., song lyrics), or task-related notes. To allow for comparison among groups, whose size ranged from 2-4 members, the total word count was adjusted for groups of 2 and 3 by multiplying by 2 or 1.33 respectively.

Similar to the independent variable, the dependent variables of accuracy and cohesion were operationalized holistically and analytically, to provide a multidimensional view of the wikis. **The holistic ratings** were performed by five German instructors at the home institution of Class 1. Their teaching experience ranged from 2-11 years and four of them had previously taught this level and this particular topic. Raters knew that these wikis were composed by students in 4th or 5th semester German. They were provided with the definition of cohesion listed below and instructed to rate the essays holistically for accuracy and cohesion on a scale of 1-5, and to briefly comment on their ratings.

Based on Halliday & Hasan (2013), the following definition of cohesion was used for the holistic ratings as well as the analytic coding: Cohesion refers to the semantic relations in a text. Cohesion is not a matter of content or textual meaning but is displayed in the linguistic ties that exist within a text (e.g., personal/demonstrative pronouns, definite articles, conjunctions).

The **analytic coding** was performed by two of the researchers, who reached an inter-rater reliability of 95%. As instructors of these courses, we relied on our knowledge of participants' proficiency in German, the programs' curricula and English native speakers' acquisition patterns to determine appropriate ways to code for accuracy and cohesion. Texts were coded for the following five cohesion markers: pronouns, definite articles, demonstratives, conjunctions and relative pronouns. These constructions represent reference and conjunction, two of the five categories of cohesive devices identified by Halliday & Hasan (2013), which we considered to be the most common cohesive markers accessible to intermediate learners of German. Counts for all five cohesion markers were combined for one overall cohesion score for each text.

In terms of accuracy, the final version of each group's wiki was coded for errors related to:

- case: accusative and nominative cases
- word order: finite verbs in 2nd position in main clause; infinite verb forms in last position in main clause; finite verbs in last position in subordinate clauses
- tense and agreement: present, simple past and present perfect tenses; separable-prefix verbs;
- personal pronouns

Since the length of groups' wiki pages varied, the error and cohesion scores were adjusted to represent 100 words and can therefore be compared across groups.

Correlations & Illustrative Cases

The first round of analysis was quantitative. To identify any relationships, correlations were run on the dependent and independent variables. Since histograms and the Schapiro-Wilk test of normality showed the data not to be normally distributed, the Spearman's rho, a non-parametric

measure of correlation, was selected.

RESULTS: COLLABORATION AS WORKLOAD

This section presents the findings for each research question based on the first operationalization of collaboration: a collaborative group is one where all members contribute their fair share to the task. The results from the quantitative analyses are presented first, followed by a discussion of illustrative cases.

Research Question 1: How many groups worked collaboratively?

Table 2 below illustrates the distribution of the workload index. There are several noteworthy results in this distribution. Almost the full range of workload is covered, with groups at the low and high end of the range as well as in the middle. Most groups, however, are in the middle. Of the 19 groups, 14 received a score between 0.67 and 1.33. In other words, there were a few groups that had either very balanced or very unbalanced workloads. In the end, only two groups, both of them pairs, had a perfectly balanced workload, and a group of 3 students displaying a very balanced workload. Results indicate no clear pattern in terms of workload due to group size, wiki topic, or class.

Table 2
Workload Index Distribution

Workload Index	**Number of Groups**	**Groups of 2**		**Groups of 3**		**Groups of 4**	
		Class 1	Class 2/3	Class 1	Class 2/3	Class 1	Class 2/3
0.33	1			1			
0.5	1						1
0.67	3			2	1		
1.0	6	1	2	1	2		
1.33	5			3	2		
1.67	1				1		
2.0	2	1	1				

Research Question 2: Is the length of the wiki texts related to the degree of collaboration within the groups?

As can be seen in Table 3, even when adjusted for group size, the length of the wiki texts varies greatly, ranging from 554 words to 2,040. The longest text was composed by a group of 2, and a group of 3 composed the shortest text. Interestingly, the only group of 4 had a relatively short text with only 655 words.

Table 3
Distribution of Total Word Counts and Workload Index

Group*	Class	Workload Index	Number of Group Members	Total Word Count	Adjusted Total Word Count
Religion	2/3	1.0	2	1020	2,040.00
Travelling	1	0.67	3	1471	1,956.43
Existentialism	2/3	1.0	2	630	1,260.00
Stasi Secret Police	1	1.0	2	594	1,188.00
Church & Religion	1	0.33	3	775	1,030.75
Housing	1	1.33	3	763	1,014.79
Potsdam Conference	2/3	0.67	3	742	986.86
Work & College	1	2.0	2	473	946.00
Sandman Kids Program	1	1.0	3	667	887.11
FDJ Youth Organization	1	1.33	3	643	855.19
FDJ Youth Organization	2/3	1.67	3	636	845.88
Youth Consecration	2/3	2.0	2	381	762.00
Socialist Unity Party	2/3	1.33	3	551	732.83
Stasi Secret Police	2/3	1.0	3	531	706.23
Banned Items	2/3	0.5	4	655	655.00
East German Products	1	1.33	3	468	622.44
Border	2/3	1.33	3	459	610.47
Banned Music	1	0.67	3	453	602.49
Iron Curtain	2/3	1.0	3	417	554.61

* Groups are identified by the topic of their wiki

With regard to the research question, a Spearman's rho showed a negative trend but was not significant: $r_s(17) = -.168$, $p = .493$. In other words, there is no statistical evidence for a relationship between text length and workload distribution.

Research Question 3: Is the quality of the wiki texts in terms of accuracy and cohesion related to the degree of collaboration within the groups?

Like above, descriptive data for the dependent variables are presented first. Table 4 below lists the number of errors and cohesion markers per 100 words, as well as the holistic ratings. The number of errors per 100 words ranged from 1.80 to 7.61 and the number of cohesion markers from 12.19 to 31.73, showing considerable variability. A similar trend is reflected in the holistic ratings, even between ratings of the same wiki page.

Table 4
Accuracy and Cohesion Ratings and Scores

Group*	Class	Number of Group Members	Errors per 100 Words	Cohesion Markers per 100 Words	Holistic Accuracy and Cohesion Ratings
Housing	1	3	2.75	12.19	4, 4, 3, 3, 2 (M=3.20)
Sandman Kids Program	1	3	1.80	20.39	5, 5, 5, 4, 4 (M=4.60)
Stasi Secret Police	1	2	3.70	22.56	4, 3, 3, 3, 2 (M=3.00)
FDJ Youth Organization	2/3	3	6.13	22.17	4, 3, 3, 3, 3 (M=3.20)
FDJ Youth Organization	1	3	3.11	31.73	5, 5, 5, 5, 4 (M=4.80)
East German Products	1	3	6.41	15.60	4, 4, 3, 3, 2 (M=3.20)
Travelling	1	3	5.10	13.60	5, 3, 3, 3, 2 (M=3.20)
Iron Curtain	2/3	3	4.56	23.02	3, 2, 2, 2, 2 (M=2.20)
Border	2/3	3	6.10	26.80	3, 3, 3, 3, 2 (M=2.80)
Existentialism	2/3	2	5.08	20.48	3, 3, 3, 3, 3 (M=3.00)
Youth Consecration	2/3	2	3.41	23.36	5, 5, 4, 3, 3 (M=4.00)
Potsdam Conference	2/3	3	2.02	20.49	5, 4, 4, 4, 2 (M=3.80)
Religion	2/3	2	2.35	14.90	5, 5, 5, 5, 4 (M=4.80)
Church & Religion	1	3	4.13	22.97	4, 4, 4, 3, 3 (M=3.60)
Socialist Unity Party	2/3	3	5.08	21.78	3, 3, 2, 2, 1 (M=2.20)
Stasi Secret Police	2/3	3	2.26	25.42	5, 4, 4, 4, 3 (M=4.00)
Work & College	1	2	7.61	21.35	5, 4, 4, 3, 3 (M=3.80)
Banned Music	1	3	6.18	18.10	4, 3, 3, 3, 2 (M=3.00)
Banned Items	2/3	4	4.58	24.43	4, 3, 3, 3, 2 (M=3.00)

* Groups are identified by the topic of their wiki

The Spearman's rho was not significant for any of the three measures: errors per 100 words ($r_s(17) = .290$, $p = .228$), cohesion markers per 100 words ($r_s(17) = .118$, $p = .630$) and holistic ratings

($r_s(17) = -.109$, $p = .656$). These analyses do not support any relationship between a group's workload distribution and the accuracy and cohesion of the text it produced.

Illustrative Cases:

Based on the results of the Spearman's rho, a more in-depth analysis of select groups was conducted. To highlight the processes and products of notably collaborative and uncollaborative groups, extreme case sampling was used and the five groups at the top and bottom ends of the workload index range were selected, based on clear cut-off values between collaborative and uncollaborative work styles. Table 5 below summarizes their scores.

The two least collaborative groups look rather similar on the accuracy and cohesion measures. Their numbers of cohesion markers are in the middle of the range (12.19-31.73), where most groups scored, and the same is true for accuracy (range: 1.80-7.61) and the holistic ratings. There is, however, a striking difference with regard to the length of their wiki texts. The group "Banned Items" had the 5th shortest text with 655 words while "Church & Religion" ranked 5th longest with 1030.75 words.

The three most collaborative groups also had rather similar numbers of cohesion markers, with their scores falling in the middle of the range, just like the least collaborative groups. While the length of their texts was fairly similar as well, they were on the short side (range: 554.61-2,040). Interestingly, one of the most collaborative groups, "Work & College", had the most errors per 100 words of all 19 groups. These findings match the results of the statistical analyses in that they fail to illustrate a connection between degree of collaboration on the one hand and text length and quality on the other.

Table 5
Profile of Illustrative Cases for Collaboration as Workload

Group		Class	Number of Members	Work-load Index	Adjusted Total of Word	Errors per 100 Words	Cohesion Markers per 100 Words	Holistic Ratings
Uncollaborative	Church & Religion	1	3	0.33	1030.75	4.129	22.968	4, 4, 4, 3, 3 (M=3.60)
	Banned Items	2/3	4	0.50	655	4.580	24.427	4, 3, 3, 3, 2 (M=3.00)
Collaborative	FDJ Youth Organization	2/3	3	1.67	845.88	6.132	22.170	4, 3, 3, 3, 3 (M=3.20)
	Youth Consecration	2/3	2	2.0	762	3.412	23.360	5, 5, 4, 3, 3 (M=4.00)
	Work & College	1	2	2.0	946	7.611	21.353	5, 4, 4, 3, 3 (M=3.80)

RESULTS: COLLABORATION AS CO-OWNERSHIP

This section focuses on the second operationalization of collaboration: a collaborative group is one where members take responsibility for the text as a whole and do not limit their edits to the sections they themselves contributed. After discussing each of the research questions in light of the quantitative analyses, we present a few illustrative cases.

Research Question 1: How many groups worked collaboratively?

Table 6 below shows the distribution of the co-ownership index for all 19 groups, which spans almost the full range of 0-100%. The members of the least collaborative group, for example, made only 3% of their edits in sections contributed by another writer (co-ownership index = 3%). The same variance can be observed when comparing the different group sizes: there are groups of 2 and groups of 3 that worked very collaboratively in terms of co-ownership, somewhat collaboratively and uncollaboratively. Results indicate no clear pattern in terms of a greater or lesser amount of co-ownership due to group size, wiki topic, or class.

Table 6
Co-ownership Index Distribution

Co-ownership Index	**Groups of 2**		**Groups of 3**		**Groups of 4**	
	Class 1	Class 2/3	Class 1	Class 2/3	Class 1	Class 2/3
3%		1				
7%		1				
10%				1		
17%				1		
20%			1			
22%				1		
25%			1			
35%		1				
42%	1					
54%				1		
56%				1		
60%			1			
65%			1			
66%			1			
68%			1			
78%						1
85%				1		
86%			1			
100%	1					

Research Question 2: Is the length of the wiki texts related to the degree of collaboration within the groups?

A Spearman's rho showed no significant correlation between the co-ownership index and the

adjusted total word count: $r_s(17) = .182$, $p = .455$. This means that how groups edited their texts was not related to their length. As Table 7 shows, there are no discernible patterns regarding length of wiki texts and co-ownership index.

Table 7
Distribution of Total Word Counts and Co-Ownership Index

Group*	Class	Co-Ownership Index	Number of Group Members	Total Word Count	Adjusted Total Word Count
Religion	2/3	35	2	1020	2,040.00
Travelling	1	68	3	1471	1,956.43
Existentialism	2/3	3	2	630	1,260.00
Stasi Secret Police	1	42	2	594	1,188.00
Church & Religion	1	86	3	775	1,030.75
Housing	1	66	3	763	1,014.79
Potsdam Conference	2/3	56	3	742	986.86
Work & College	1	100	2	473	946.00
Sandman Kids Program	1	65	3	667	887.11
FDJ Youth Organization	1	25	3	643	855.19
FDJ Youth Organization	2/3	22	3	636	845.88
Youth Consecration	2/3	7	2	381	762.00
Socialist Unity Party	2/3	17	3	551	732.83
Stasi Secret Police	2/3	85	3	531	706.23
Banned Items	2/3	78	4	655	655.00
East German Products	1	60	3	468	622.44
Border	2/3	54	3	459	610.47
Banned Music	1	20	3	453	602.49
Iron Curtain	2/3	10	3	417	554.61

* Groups are identified by the topic of their wiki

Research Question 3: Is the quality of the wiki texts in terms of accuracy and cohesion related to the degree of collaboration within the groups?

Just like all the other correlations, there were no significant results based on co-ownership for the analytic error ($r_s(17) = -.060$, $p = .808$) and cohesion scores ($r_s(17) = -.065$, $p = .792$) on the one hand and the holistic ratings ($r_s(17) = -.303$, $p = .207$) on the other. Therefore, group members' edits of others' contributions were not related to the quality of the text.

Illustrative Cases:

Like above, extreme case sampling was used to identify the most and least collaborative groups in terms of co-ownership. Based on clear breaks in the distribution of the co-ownership index, three groups each were selected from the top and bottom end of the range. The quantitative data for those six groups is shown in Table 8 below.

Table 8
Profile of Illustrative Cases for Collaboration as Co-ownership

Group	Class	Number of Members	Co-ownership Index	Adjusted Total of Word	Errors per 100 Words	Cohesion Markers per 100 Words	Holistic Ratings
Existentialism	2/3	2	3	1260	5.080	20.476	3, 3, 3, 3, 3 (M=3.00)
Youth Consecration	2/3	2	7	762	3.412	23.360	5, 5, 4, 3, 3 (M=4.00)
Iron Curtain	2/3	3	10	554.61	4.556	23.022	3, 2, 2, 2, 2 (M=2.20)
Stasi Secret Police	2/3	3	85	706.23	2.260	24.424	5, 4, 4, 4, 3 (M=4.00)
Church & Religion	1	3	86	1030.75	4.129	22.968	4, 4, 4, 3, 3 (M=3.60)
Work & College	1	2	100	946.0	7.611	21.353	5, 4, 4, 3, 3 (M=3.80)

Comparing the most with the least collaborative groups in terms of co-ownership, several similarities were found. There is a mix of groups of 2 and 3 and their cohesion marker numbers are rather analogous, representing the middle of the range (range: 12.19-31.75). However, the third least collaborative group, "Iron Curtain", produced the shortest text of all the 19 groups and also received the lowest mean rating from the five raters. Interestingly, the overall poor quality that 4 of the 5 raters perceived is not reflected in the analytic error and cohesion marker numbers, which are average compared to other groups. The following rater comments provide possible reasons for this mismatch:

- *"Simple sentence structure … Occasional use of conjunctions."*
- *"Mostly a succession of main clauses; if at all, only implementation of basic conjunctions like 'and' and 'because'."*
- *"Plenty of mistakes which often affect the understanding of the text; at times total lack of cohesion."*
- *"Some weird word order that makes it difficult to follow."*

These comments illustrate that the analytic scores do not capture the texts' lack of complexity and the effect of errors and lack of cohesion on the reader. A similar mismatch occurred with the "Youth Consecration" group: while its analytic cohesion score was almost identical to the previously discussed group, "Iron Curtain", the "Youth Consecration" group's holistic rating ranked 4th in all ratings, which was also supported by the raters' favorable comments:

- *"Cohesion: Good use of conjunctions such as 'because' and 'but' and concepts are referred to with personal pronouns in following sentences, making the ideas clear and the text logically structured."*
- "*Excellent cohesion through command of conjunctions and temporal markers; seamless*

transition between paragraphs."

Looking at the most collaborative groups at the opposite end of the spectrum, it is interesting that the "Stasi" group had a rather low occurrence of errors (2.260; range: 1.8-7.61) while the "Work & College" group actually produced the text with the most errors overall. Ultimately, instructors hope that learners will edit each other's work and do so successfully. The example of these two groups, whose members revised the contributions of others, shows that this is not always the case.

The "Work & College" pair is a particularly interesting case, as its co-ownership index of 100% means that all of the edits were made in sections contributed by the partner. While this seems extreme at first glance, a closer look at their archived edits shows that both members made a rather low number of edits, 3 and 4 respectively, in the areas of spelling, punctuation, word order, and vocabulary. The low number of edits is, however, not uncommon, as there were seven other groups in which one or more members made less than five edits in someone else's text. It is also important to acknowledge that the wiki's archived versions do not capture the edits that writers make until the page is saved for the first time. It is therefore likely that these two writers edited as they wrote their initial versions, rather than not editing their own sections at all.

Raters' comments on the two most collaborative groups in terms of co-ownership, "Work & College" and "Church & Religion", provided further interesting insights. Although the group "Work & College" made all of their edits in sections originally written by another group member, raters noticed differences between the first paragraph and the following sections:

- *"Accuracy: strong differences between the first and the second part. Very good structures in the first part, problems with verb forms and word order mostly in the second part. Cohesion: There are again a few differences between both parts."*
- *"The first part of this is definitely a five, but there are quite a few mistakes in the second one."*
- *"Very cohesive text in the first half; less so in part two."*

Correspondingly, the wiki page of the "Church & Religion" group, which made 86% of its archived edits on others' contributions, led raters to make the following similar observations:

- *"Starting from the second paragraph, the writing is almost perfect. The first paragraph, however, was a bit confusing because the level of complexity the writer was trying to achieve was too high for him/ her."*
- *"Accuracy: Good use of various structures after the first paragraph."*

While discrepancies between different parts of a wiki text can be interpreted as a sign that other group members did not edit each other's parts, that was not the case in these two groups. It further shows that holistic ratings capture nuances of a text that analytic scores cannot.

In addition to illustrating the value of combining analytic analyses with a more holistic, qualitative assessment, these illustrative cases show the lack of any relationship between the

dependent and independent variables.

DISCUSSION

This study explored the connection between the process of collaborative wiki writing and the length, accuracy and cohesion of the produced texts. Based on the proposition that true collaborative writing involves both a shared workload and joint responsibility for the text, collaboration was operationalized by workload distribution and the amount of edits members made in segments that were originally written by other group members. Both of them will be compared and discussed here in connection to other studies.

The first research question examined how many groups worked collaboratively. For both operationalizations, results illustrated that few groups worked in a truly collaborative manner. In terms of workload, only three groups were very balanced, two groups showed a very unbalanced workload, and the majority of the groups (14) ranged in the middle, revealing a somewhat unbalanced workload. These unbalanced workloads might be attributed to: 1) students' different levels of task motivation, including both intrinsic and extrinsic factors (Arnold, Ducate, & Kost, 2012; Kessler & Bikowski, 2010), 2) unfamiliarity with the online medium, 3) uneasiness with assumed roles or with writing in a foreign language as observed in Kessler, Bikowski, and Boggs' (2012) study, or 4) little experience with or dislike for this kind of task.

In terms of co-ownership, the 19 groups showed a very even distribution across the full span of the co-ownership index, ranging from one group that made only 3% of their edits in another member's original writings, to one group that exclusively edited each other's work. Interestingly, one group, "Work & College", had both the highest workload and co-ownership indexes, meaning that it met both criteria for true collaboration. However, this group produced the least accurate text of all 19 groups with a medium cohesion score. As discussed above, though, the two members made all of their archived edits in sections originally written by the partner, which might not be ideal.

Overall, edits in other members' texts varied widely across the groups, ranging from 1-196 edits, with an average of 20.5 edits, and a median of 9. This wide spread mirrors findings from previous studies in which participants either seemed reluctant to edit others' contributions because they did not want to threaten face or to trust their own language skills (Lee, 2010; Mak & Coniam, 2008) or edited others' texts to the same or even larger extent than their own (Kessler, 2009). Organized feedback from peers and instructors has also been shown to lead to more edits in participants' own writings than in their group members' texts (Arnold, Ducate, & Kost, 2012). The divergent findings further underline the value of task communication, which can occur face-to-face (Martinsen & Miller, 2012) or via computer-mediated communication (Elola & Oskoz, 2010; Yeh, 2014; Zheng, Niiya & Warschauer, 2015), and of learner training for peer editing to decrease apprehension when editing and to promote a deeper feeling of co-ownership of a produced text (Min, 2006).

The second research question explored whether the length of the wiki texts was related to the degree of collaboration in the groups. Although the adjusted length of the texts differed greatly,

ranging from 554 to 2040 words, neither operationalization showed a correlation. In terms of workload, the three groups with the most collaboration are in the middle range as far as text length is concerned, whereas the two groups with the least are in the top and bottom quarter of the list, respectively. When looking at text length from the perspective of co-ownership, uncollaborative groups produced some of the shortest, medium-length, as well as one of the longest texts, and very collaborative groups showed a similar variance.

When considering group size, it seems that pairs wrote longer texts than larger groups as the dyads are located in the upper half of the word count list. These findings could indicate that groups with fewer members wrote longer texts, similar to Storch's (2005) findings. It has to be kept in mind, though, that in her study, participants worked together in an offline, paper-based classroom setting which might further underscore limited comparability between asynchronous collaborative writing on wikis and synchronous collaborative writing in a face-to-face format. Results of the present study, however, are in contrast to the findings in Strobl's (2014) study, in which groups of three members wrote longer texts than individuals, which might have been due to a more cooperative writing style of assembling separate pieces for the final product or their higher proficiency level. However, comparisons with previous research are difficult due to differences with regard to group size, task type and the way collaborative writing was defined.

The third research question investigated whether accuracy and cohesion of the wiki texts was related to the degree of collaboration within the groups. Again, statistical analyses revealed no correlation between the groups' accuracy and cohesion scores on the one hand and their workload distribution and co-ownership of the text on the other. In order to gain a multidimensional perspective of accuracy and cohesion, the wiki texts were also rated holistically, but these ratings again showed no pattern. A closer look at the illustrative cases confirms the large variability and the lack of connections between text length, text quality, and workload or co-ownership. This corroborates findings in other studies where no difference was found between collaborative wiki texts and individual texts in terms of complexity, accuracy, fluency, coherence and cohesion (Elola & Oskoz, 2010; Strobl, 2014). Strobl (2014), who worked with higher proficiency students, was somewhat surprised "that the individual texts do not achieve higher levels of cohesion and coherence than the collaborative texts. After all, the groups had to agglutinate three individually written pieces, which could have resulted in a less coherent joint text." (p. 11). The same is true for the current project, in which groups with an unequal distribution of workload produced texts that were just as cohesive as the ones produced by groups with a more balanced workload. Correspondingly, groups that made a larger number of edits in texts originally written by other group members produced similarly cohesive texts as those groups that made less edits.

While Elola and Oskoz (2010) also could not detect any statistical differences in fluency, accuracy, and complexity between individual and collaborative writings, they felt that the collaborative writings were more successful as the groups were able to express their ideas more fully. This seems to mirror findings in the current study where holistic ratings based on accuracy and cohesion of the texts did not necessarily reflect quantitative measures. For example, in the current study, of the two groups with the highest holistic rating of 4.8, "Religion" had the third lowest cohesion score of 14.9, whereas "FDJ Youth Organization" had the highest cohesion score of 31.73. The discrepancy between analytic measures and holistic ratings was also evident in the

groups exhibiting the lowest holistic scores of 2.2, "Iron Curtain" and "Socialist Unity Party", which both received cohesion scores in the middle of the range. The mismatch between statistical results, accuracy and cohesion scores, and holistic ratings were further underscored by the raters' comments, which showed that the scores do not necessarily align with a reader's reception of a text. It is possible that raters might have been influenced by the content and coherence of a text, even though they were instructed to base their ratings solely on accuracy and cohesion.

One reason for the observed mismatches might be that even instructors who are used to analytic essay rubrics can find it difficult to separate all dimensions completely and clearly from one another. Features of one dimension might inadvertently affect instructors' judgment of another. As documented in the assessment literature (Hughes, 2003), analytic scoring can fail to capture that writing is more than the sum of its parts. Another reason for the mismatch between analytic accuracy and cohesion scores and holistic ratings might be the nature of the edits that were made by the students. The revisions that were examined were restricted to changes intended to increase the grammatical accuracy of a text, and to those that intermediate L2 learners are usually able to detect and correct, such as spelling, word order, vocabulary, verb endings, and cases, but not to cohesive devices. While it was expected that basic cohesive markers, such as possessive pronouns and coordinating conjunctions, could be identified and edited by intermediated learners, they might need more guidance and training in regard to these structures' role in cohesion, in addition to more advanced markers, including conjunctive elements such as adverbs (e.g., therefore) and prepositional phrases (e.g., in contrast).

CONCLUSION

In conclusion, the present study demonstrates that many groups did not truly collaborate and that the degree of collaboration did not affect the cohesion, accuracy and length of the product. These findings have several pedagogical implications. The fact that an uneven workload was characteristic of many groups underscores the importance of student accountability. This can be achieved by using a jigsaw format or collaborative tasks (Aydin & Yildiz; 2014; Lee, 2010; Lund, 2008), assigning roles (Yeh, 2014; Zheng, Niiya, & Warschauer, 2015), increasing task visibility (i.e., knowing that the instructor is aware of each student's efforts; Piezon & Donaldson, 2005), creating weekly milestones and pre-, during-, and post- writing activities to structure progress and reflection throughout the process (Alyousef & Picard, 2011; Limbu & Markauskaite, 2015), and using rubrics for peer evaluation (Alyousef & Picard, 2011; Zheng, Niiya, & Warschauer, 2015). Furthermore, this finding reminds us of the central role of task motivation, which can be fostered with relevant enjoyable tasks that allow learners to experience success and exercise agency and autonomy (e.g., by selecting the topic) (Guilloteaux & Dörnyei, 2008). Participants' general preference for self-editing might be a sign that they were either reluctant to work on what they perceived as "someone else's text" or that they did not see the value of peer-editing. In addition to pre-task training, these concerns can be addressed by encouraging groups to use the various collaboration tools available in online writing environments, such as a discussion board in a wiki (Zheng, Niiya, & Warschauer, 2015).

Somewhat surprisingly, the present study did not show that collaboration affected the quantity

or quality of the texts the groups produced. At first, this seems to dispute recent collaborative methodologies; however, it could be that the analyses used here were unable to uncover advantages of collaboration or that those benefits might appear in subsequent tasks or in other contexts. This points toward several limitations of this study. The current study focused on only one large writing task and did not include any complexity or content measures, which would provide further insights into text quality. Furthermore, the present study did not control for group size but found indications that it might be a powerful extraneous variable. By analyzing the wikis' archived version, it also did not capture any edits learners made in their own writing before saving the wiki text. In addition, when assigning holistic ratings to the texts, raters might have been influenced by a text's content or coherence even though they were instructed to only focus on issues of accuracy and cohesion. Lastly, as non-experimental classroom research, this study did not yield any generalizable results.

These limitations can inform future research, which could examine students' edits and final products more qualitatively. For example, is there a difference in the type of edits that learners make in their own vs. others' contributions? Are there any patterns in terms of whose contributions they edit? Do the process and outcomes of group writing tasks change over time as learners become more familiar with each other and the medium? It would also be interesting to investigate learners' editing processes (e.g., similar to Smith, 2008), to shed light on how and why learners make changes to the wiki. Further studies could also examine the effect of training on students' success at collaboration, a recommendation that has been included in many studies. As this study and other research have shown, there are still many angles from which to examine group writing. The current study offers some insights into the importance of defining the term collaboration in research and the classroom and emphasizes the need to further investigate if and how task processes affect task outcomes.

ABOUT THE AUTHORS

Nike Arnold (Ph.D., University of Texas at Austin) is Associate Professor of Applied Linguistics at Portland State University in Oregon, USA. She teaches a variety of language pedagogy courses and her research focuses on educational technology as well as social and affective aspects of learning and teaching another language.

Lara Ducate (Ph.D., University of Texas at Austin) is a Professor of German and Applied Linguistics at the University of South Carolina. Her research interests include teacher education, intercultural learning, computer assisted language learning, and study abroad.

Claudia Kost (PhD University of Arizona) is Associate Professor of German Applied Linguistics at the University of Alberta, Canada. Her research interests focus on wikis, vocabulary acquisition, and pedagogical implications for the foreign language classroom.

REFERENCES

Allen, N., Atkinson, D., Morgan, M., Moore, T., & Snow, C. (1987). What experienced collaborators say about collaborative writing. *Journal of Business and Technical Communication, 1*(2), 70-90.

Alyousef, H. S., & Picard, M.Y. (2011). Cooperative or collaborative literacy practices: Mapping metadiscourse in a business students' wiki group project. *Australasian Journal of Educational Technology, 27*(3), 463-480.

American Council on the Teaching of Foreign Languages. (2012). *ACTFL Proficiency Guidelines*. Alexandria, VA: ACTFL.

Amir, Z., Ismail, K., & Hussin, S. (2011). Blogs in language learning: Maximizing students' collaborative writing. *Procedia Social and Behavioral Sciences, 18*, 537-543.

Arnold, N., Ducate, L., & Kost, C. (2009). Collaborative writing in wikis: Insights from culture projects in German classes. In L. Lomicka and G. Lord (Eds.), *The next generation: Social networking and online collaboration in foreign language learning* (pp. 115-144). San Marcos, TX: CALICO.

Arnold, N., Ducate, L., & Kost, C. (2012). Collaboration or cooperation? Analyzing group dynamics and revision processes in wikis. *CALICO Journal, (29)*3, 431-448.

Aydin, Z., & Yildiz, S. (2014). Using wikis to promote collaborative EFL writing. *Language Learning & Technology, 18*(1), 160-180.

Brussig, T. (2003). *Am kürzeren Ende der Sonnenallee*. Stuttgart, Germany: Klett-Cotta.

Dillenbourg P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1-19). Oxford: Elsevier.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In E. Spade & P. Reiman (Eds.), *Learning in humans and machine: Towards an interdisciplinary learning science* (pp. 189-211). Oxford, UK: Elsevier.

Elola, I., & Oskoz, A. (2010). Collaborative writing: Fostering foreign language and writing conventions development. *Language Learning & Technology, 14*(3), 51-71. Retrieved from http://llt.msu.edu/vol14num3/elolaoskoz.pdf

Guilloteaux, M.J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly, 42*(1), 55-77.

Halliday, M.A.K., & Hasan, R. (2013). *Cohesion in English*. New York: Routledge.

Haythornthwaite, C. (2006). Facilitating collaboration in online learning. *Journal of Asynchronous Learning Networks, 10,* 1-24. Retrieved from http://sloanconsortium.org/sites /default/files/v10n1_2haythornthwaite_0.pdf.

Hughes, A. (2003). *Testing for language teachers.* Cambridge, UK: Cambridge University Press.

Kessler, G. (2009). Student-initiated focus to form in wiki-based collaborative writing. *Language Learning & Technology, 13*(1), 79-95. Retrieved from http://llt.msu.edu/vol13num1/kessler.pdf.

Kessler, G., & Bikowski, D. (2010) Developing collaborative autonomous learning abilities in computer mediated language learning: Attention to meaning among students in wiki space. *Computer Assisted Language Learning, 23*(1), 41-58.

Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning & Technology, 16*(1), 91-109.

Lee, L. (2010). Exploring wiki-mediated collaborative writing: A case study in an elementary Spanish course. *CALICO Journal, 27*(2), 260-276.

Limbu, L., & Markauskaite, L. (2015). How do learners experience joint writing: University students' conceptions of online collaborative writing tasks and environments. *Computers & Education, 82,* 393-408.

Lin, O. P., & Maarof, N. (2013). Collaborative writing in summary writing: Student perceptions and problems. *Procedia – Social and Behavioral Sciences, 90,* 599-606.

Lowry, P., Curtis, A., & Lowry, M. (2004). A taxonomy of collaborative writing to improve empirical research, writing practice, and tool development. *Journal of Business Communication, 41*(1), 66-99.

Lund, A. (2008). Wikis: a collective approach to language production. *ReCALL, 20,* 35-54.

Mak, B., & Coniam, D. (2008). Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System, 36*(3), 437-455.

Martinsen, R., & Miller, A. (2012). Collaboration through wiki and paper compositions in foreign language classes. *IALLT Journal of Language Learning Technologies, 42*(1), 72-95.

Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15,* 118-141.

Oskoz A., & Elola I. (2012). Understanding the impact of social tools in the FL writing classroom: Activity theory at work. In Kessler, G., Oskoz, A. & Elola, I. (Eds.) Technology across writing contexts and tasks, pp. 131-154. CALICO Monograph Series, volume 10, CALICO: San Marcos, Texas.

Piezon, S., & Donaldson, R. (2005). Online groups and social loafing: Understanding student-group interactions. *Online Journal of Distance Learning Administration, VIII, #IV*. Retrieved from http://www.westga.edu/~distance/ojdla/winter84/piezon84.htm.

Shehadeh, A. (2012). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing, 20*(4) 286-305.

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning and Technology, 12*(1), 85-103.

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System, 27*(3), 363-374.

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*, 153-173.

Strobl, C. (2014). Affordances of Web 2.0 technologies for collaborative advanced writing in a foreign language. *CALICO Journal, (31)*1, 1-18.

Wigglesworth, G., & Storch, N. (2009). Pairs versus individual writing: Effects on fluency, complexity, and accuracy. *Language Testing, 26*(3), 445-466.

Xuanxi, L., Chu, S., & Ki, W. W. (2014). The effects of a wiki-based collaborative process writing pedagogy on writing ability and attitudes among upper primary school students in Mainland China. *Computers & Education, 77*, 151-169.

Yang, L. (2014). Examining the mediational means in collaborative writing: Case studies of undergraduate ESL students in business courses. *Journal of Second Language Writing, 23*, 74–89.

Yeh, H. C. (2014). Exploring how collaborative dialogues facilitate synchronous collaborative writing. *Language Learning & Technology*, 18(1), 23–37.

Zheng, B., Niiya, M., & Warschauer, M. (2015). Wikis and collaborative learning in higher education. *Technology, Pedagogy and Education*, 1-18.