Synchronous Writing in the Classroom: Undergraduates' Collaborative Practices and their Impact on Text Quality, Quantity, and Style

Soobin Yim
UC Irvine
Irvine, U.S.
soobiny@uci.edu

Dakuo Wang
UC Irvine
Irvine, U.S.
dakuow1@uci.edu

Judith Olson
UC Irvine
Irvine, U.S.
jsolson@uci.edu

Viet Vu
UC Irvine
Irvine, U.S.
vqvu@uci.edu

Mark Warschauer
UC Irvine
Irvine, U.S.
markw@uci.edu

ABSTRACT
Group activities that use Google Docs for simultaneous collaborative writing and editing are increasingly common in higher education. Although studies show that synchronous collaboration can bring multiple benefits, such as enhanced productivity and writing quality, little is known about these writing practices in classrooms and their impact on students' writing. Using a mixed method approach, we conducted an empirical study that explores the different styles of synchronous collaboration in 45 Google Docs documents produced by 82 undergraduate students, and how students' practices affect the specific dimensions of the final text including quality. The results suggest that (a) out of four styles, *Divide and Conquer* style tended to produce better quality text whereas *Main Writer* had the lowest quality scores, and that (b) balanced participation and amount of peer editing led to longer texts with higher quality scores for content, evidence, but not organization or mechanics. Given these results, we suggest several design features for collaborative writing systems and propose guidelines for instructional practices.

Author Keywords
Collaborative writing; cloud-based technology; higher education; Google Docs; information visualization; computational text analysis; synchronous collaboration; collocated; computer-supported cooperative work.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

CSCW '17, February 25-March 01, 2017, Portland, OR, USA
 ISBN 978-1-4503-4335-0/17/03...$15.00
DOI: http://dx.doi.org/10.1145/2998181.2998356

INTRODUCTION
Collaborative writing, which refers to the joint production or co-authoring of a document where writers share in the ownership of a text [31], is increasingly common today, particularly with the wide availability of web-based tools. For the past 30 years, researchers in the field of HCI have studied how to build collaborative writing systems to support co-authors (e.g., [21]), and how users use these tools [24]. Drawing from interviews, observations, and experiments, researchers have found fruitful empirical results as well as theoretical frameworks. Among those findings, Dourish and Bellotti [10] highlighted the value of collaboration awareness, that co-authors need to know who is doing what, when, and where. As for the theoretical frameworks, Posner and Baecker [26] proposed a seminal framework to study collaborative writing, which we will describe in more detail in the related work section. Recently, cloud-based technology that enables simultaneous writing has significantly transformed the level and scope of collaborative work. Several studies in HCI have examined the new ways of collaboration in Google Docs, and have discussed topics such as the relationship between user perceptions and editing behaviors [4; 5], and work styles in a corporate environment [30].

In parallel, researchers from the field of education have studied the use of collaborative writing for learning. Collaborative group work is increasingly popular in classrooms as the enhanced sharing features available in cloud-based technology supports effective instruction. Drawing from sociocultural theories of learning, research has found that collaborative writing enhances writing quality [31], sense of audience [29], the pooling of knowledge and ideas [9], and opportunities to socialize with specific discourse communities [36]. Most of the previous studies examined the practices of collaborative writing in an asynchronous mode. For example, Kessler and Bikowski [16] focused on asynchronous feedback and commenting practices and found that collaboration typically occurred at the later stage of writing, rather than throughout the writing process. However, little is known about the ways students collaborate when they can write in a synchronous, collocated environment, such as with Google Docs, and how

various aspects of such collaboration (e.g., number of co-authors, participation equity) may relate to a document's quality and quantity.

In this paper, we aim to address these topics by analyzing 45 documents created by 82 undergraduate students working synchronously and collocated in 15 groups at various times in an academic quarter. Using a mixed method approach that incorporate information visualization (InfoVis), Natural Language Processing (NLP) text analysis methods, and rubric-guided grading of the essays by two coders, we aim to understand:

- what the different styles of synchronous, collocated collaborative writing are in this setting,
- how various characteristics of collaborative writing practices relate to the document quality and quantity, and
- what implications we can draw for both system design and instructional practices.

To obtain a thorough understanding of the resulting document, our analysis examines the final texts at multiple linguistic levels: textual (e.g., organization, textual cohesion), syntactic (e.g., syntactic complexity), and lexical level (e.g., lexical sophistication). The results help us examine whether collocated, synchronous collaborative writing practices result in high quality outcomes. We further propose both design and educational implications for integrating and improving collaborative writing technologies in higher education.

RELATED WORK

Collaborative Writing Research in HCI

From the early days, researchers found that people wrote together because collaboration improves work efficiency and sometimes the quality of an outcome [15; 28]. People wrote various types of documents together such as business reports and homework assignments [11]. Seeing that people have the need for collaborative writing, HCI pioneers from various disciplines such as psychology, computer science, and engineering, built and tested various experimental tools with features to support collaborative writing. One of these experimental systems, SASSE, provided a unique feature of explicit role assignment (e.g., writer and reviewer), an enhanced communication channel, and features that magnified collaboration awareness between users [3]. Later, Mitchell et al. [23] designed an experiment in which two groups of four students in the sixth grade used SASSE. The researchers found that even though the students had never written collaboratively or had prior experience with the tool, they managed to use the synchronous editing feature and finished the writing task. During the process, these participants also developed the sense of collaboration awareness, ownership, and new control strategies to adopt synchronous editing.

ShrEdit was another experimental system that was built in the 1990s [21]. As shown in Figure 1, ShrEdit supported a

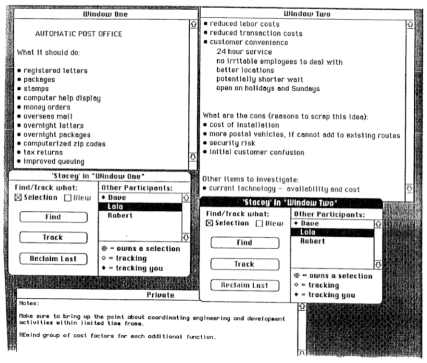

Figure 1. The user interface of ShrEdit [24].

private editing window together with co-authoring windows. It also allowed co-authors to edit each other's work simultaneously at the keystroke level, a feature that is now available on Google Docs. Since ShrEdit was mainly designed for collocated writing, it did not provide an explicit communication channel.

HCI pioneers not only built systems but also evaluated these systems using many research methods adopted from other fields. For example, Olson et al. [24] designed an experiment to compare how groups wrote together with traditional technology (i.e., whiteboard, pen and paper) and with the new shared editor (i.e., ShrEdit). They found that the groups with ShrEdit generated fewer but better design ideas. The authors hypothesized that this was likely because there was a shared focus on the collaboratively written documents.

Posner and Baecker [26] synthesized other researchers' empirical findings with their interviews of ten people reporting about 22 collaborative writing projects. From the interview data, they answered a number of questions about writing collaboration:

- what are users' expectations about collaborative writing;
- what are the social dynamics (e.g., authorship and trust);
- what are the technologies the users used to write and to communicate; and
- what writing strategies do users use.

The authors then developed a theoretical framework to describe various aspects of collaborative writing, as summarized in Table 1.

Posner and Baecker noted that collaborative writing had six distinct **activities**: brainstorm, research, plan (for the content and for the process), write, edit, and review. They noted that people played different **roles**: writer, consultant, editor, reviewer, and equal work. They also described **document control methods** and **writing strategies** (i.e., style of work). For example, horizontal division, which members can also call divide-and-conquer, requires the group to plan and divide the work, with each of the co-authors writing a section, allowing them to work in parallel.

Role	Activities	Document control methods
Writer Consultant Editor Reviewer Equal work	Brainstorm Research Plan Write Edit Review	Centralized Relay Independent Shared
Writing strategy		**Work mode**
Single author (i.e., Scribe) Horizontal division Reactive writing Parallel writing Sequential writing		Degree of proximity Synchronicity Mixed mode

Table 1. Posner and Baecker's [26] framework of collaborative writing, extended by Lowry et al. [19].

Lowry et al. revisited the research on collaborative writing using the newly-emerged, Internet-based, distributed tools from the 2000s [19]. Their primary goal was to provide a taxonomy to unite the researchers and practitioners who were studying collaborative writing from different disciplines (e.g., computer science, social science, and humanities). They extended and modified Posner and Baecker's framework to include work modes, which refers to "when and where the group will do its writing, in terms of same or different place and same or different times," as included in Table 1.

Synchronicity and Distance in Collaborative Writing

Some research focused on a few specific aspects of collaborative writing, such as the synchronicity and the distance, to see how these aspects interacted with group dynamics. Birnholtz et al. [5] designed a laboratory study in which two people worked together to write a document. They focused on how co-authors communicate while they are writing in Google Docs, and whether this communication impacts their social relationship. They also hypothesized that the direct editing of another's text would negatively impact their social relationship. The results showed that communication (e.g., real-time text interaction, commenting) helps to maintain people's social relationships in synchronous writing, but harms social relationships in asynchronous writings. One explanation was that because asynchronous writing does not require as much communication as synchronous writing (synchronous writing requires coordination to avoid collisions), when there was a lot of communication, coauthors interpreted it as wasting time. The number of edits also has a negative impact on people's feelings about each other. Therefore, they suggested that Google Docs should use "<Name> suggests <changes>" instead of showing the changes directly, which is reflected in the "suggesting" feature introduced in Google Docs in 2015.

Collaborative Writing Research in Higher Education

With the maturation of the Internet and the growing penetration rate of computers and high speed internet access in schools and homes, educators and researchers are turning their attention to online technologies such as Google Docs to enhance collaboration, facilitate communication, and share information [6; 14]. To verify the educational benefits of collaborative writing practices, several studies have examined the impact of technology-based collaboration on student learning, particularly on their writing. For example, Mak and Coniam [20] examined the textual quality of wiki-based collaborative products. The authors used both descriptive textual analysis (i.e., descriptions of changes in textual measures across multiple drafts) and qualitative analysis, and found improvements in text quality (e.g., complexity, coherence) and quantity from a previous draft. They suggested that these improvements are attributable to the collaborative nature of the task and the strong presence of an audience. Although the approach is valuable, the limited sample size and the lack of assessment of the quality of final writing outcome make it difficult to draw a reliable conclusion about how this may lead to writing improvement.

Other experimental studies compared the differences between individual and collaborative writing products (e.g., [1; 32]). These studies examined the impact of collaborative writing on texts of second language writers. The findings from these studies suggested that collaborative writing is useful for improving content and organization. For example, Arslan and Şahin-Kızıl [1] examined the effect of blog-based writing instruction on students' writing performance. Compared to the control group, the blog intervention group showed greater improvements in their writing, particularly in content and organization, but not in other areas such as vocabulary and grammar. Other studies rendered support to these findings (e.g., [35]), suggesting that writing in collaborative online environments may promote students' awareness of readers and may contribute to a more organized and clearer message.

More recently, Google Docs has been increasingly popular due to its enhanced sharing feature that enables synchronous group writing and editing. Research on synchronous collaborative writing in higher education settings is only emerging and there are only a handful of studies. For example, Yeh [37] analyzed how students' collaborative discussion impacts textual quality and found that highly collaborative groups (i.e., groups that engage in more discussion) produced better quality essays in terms of fluency and accuracy. The author suggested that during synchronous collaborative writing, students are exposed to rich linguistic input, which can contribute to writing development.

However, there is little empirical evidence on how students incorporate synchronous writing in their classroom group assignments in naturalistic settings, and whether the apparent benefits of asynchronous collaboration on improved content and organization of text (e.g., [1]) would

transfer to synchronous collaborative writing. It is also unclear what factors of collaboration contribute to better writing, and which dimensions of writing (e.g., content, organization) are being affected.

Answering this question has been difficult partly because we lack quantifiable data about collaborative behavior. In contrast to most previous education studies that were qualitative case studies with observations and interviews, we use a tool that offers both information visualization and calculations of editing amount and participation equality. In measuring the text quality, we used both human grading based on an analytic rubric and Coh-Metrix [8], a computational text analysis tool that utilizes natural language processing techniques. This tool provides over 200 indices of textual features that reflect cohesion relations, word knowledge, and language and discourse characteristics [13] through modules such as syntactic parsers and latent semantic analysis. Studies have validated the tool's predictive capacity for measuring textual difficulty and readability [8; 22]. By utilizing these tools, this study attempts to reveal new insight on how collaborative writing behaviors may impact specific textual characteristics in writing.

Research Questions

Prior studies have enriched our understanding of writing collaborations, but we lack a detailed understanding of students' synchronous, collocated collaborative writing practices in a classroom setting. This investigation is important given the popularity of synchronous group work in education and its potential as a promising new technique for instruction. In this paper, we aim to investigate students' various ways of synchronous writing and their writing outcomes. We are also interested in how the diverse ways of writing may relate to a specific, detailed set of outcomes. Thus, we propose the following two research questions:

RQ1: What are the undergraduates' different styles of writing when they write together in a synchronous, collocated classroom setting?

RQ2: How do the characteristics of collaboration (i.e., participation equality, editing amount, editing source) relate to text quality (i.e., rubric-guided quality, computational traits) and quantity?

First, we are interested in how student groups write in the synchronous mode (i.e., practices), and what characterizes the different practices. We would intuitively expect that the ability to write simultaneously together in one location might manifest new practices of collaborative writing. However, it is also plausible that students might maintain the old ways of working together and might not utilize the affordances of technology or take advantage of the work environment. In this context, the students might decide to take turns to write in a document (Sequential Writing in Posner & Baecker's taxonomy [26]), or delegate one student to write the whole document while others only provide comments and ideas (Scribe in [26]). In addition, we examined how the writing styles differ in terms of collaboration behaviors (e.g., participation equality, group activeness), text quality, and quantity.

Second, we aim to explore the effects of synchronous writing practices on document quality. Previous studies on collaborative writing in an asynchronous mode suggest that collaboration enhances content and organization [1], yet we have little empirical evidence about whether these benefits may be applicable to synchronous mode of collaboration, and if so, which aspect of collaboration may contribute to the result. In addition to examining the writing quality at the document level (e.g., content, organization), as typically done in prior studies [1; 32], we utilize a computational text analysis software (i.e., Coh-Metrix) to understand the impact of collaboration on specific traits at multiple linguistic levels.

Using the linguistic trait measures (e.g., lexical diversity, syntactic complexity), we test, for example, if student groups use more diverse types of vocabulary and produce syntactically complex sentences when members participate equally. With the support of computational text analysis, we also analyze students' text quantity on multiple levels: word count, sentence count, and paragraph numbers and length.

Using the quantity outcomes, we could test, for example, whether groups that exhibit balanced participation produce longer documents compared to those with a few main writers, and if so, at which linguistic level. Given that the amount of writing produced in a timed setting (i.e., 50 minutes in our study) is often used as a measure of writing fluency [18], the results may imply how collaboration characteristics may impact writing fluency and team productivity. Finally, using the various text outcome measures, we examined whether text quantity and quality are related to each other.

Our ultimate goal is to produce design implications for collaborative writing system designers, and to suggest guidelines for students and instructors who are participating in collaborative writing activities in higher education. How can we design systems to better support students' collaborative writing activities in a synchronous, collocated classroom setting? We want students to benefit from group writing; we want to facilitate classroom instruction that easily coordinates and manages student collaboration; and we want to support student teams to produce higher quality and quantity outcomes in a shorter time.

RESEARCH METHOD

Our data corpus consists of 45 documents written via Google Docs and their fine-grained data traces (at millisecond level and at keystroke level). These documents were written by 15 groups of undergraduate students who stayed in the same group throughout the quarter (10 weeks). The 82 students were randomly assigned to groups, with one group of four students, six groups of five students and eight

groups of six students. The students were in three sections of an undergraduate-level course offered by the School of Education at a large university in the United States. A majority of the students were pursuing a degree in Education (72%) and they were from diverse academic years (Freshmen 35.1%, Sophomores 16.2%, Juniors 20.3%, and Seniors 28.4%).

During class, student groups wrote synchronously three persuasive essays on educational themes, producing one essay per week for three weeks. For each source-based writing assignment, student groups did individual research about the topic before class, and were allowed 50 minutes in class to share their sources, plan, discuss, and write a group essay that addressed an instructor-developed prompt related to the course's weekly topic (e.g., how to best support digital literacy in the K-12 classroom). At the beginning of the course, the instructor gave a brief introduction to the collaborative features of Google Docs, but did not give specific instructions on how to collaborate.

We have the final documents produced and the traces of who wrote what when (with their consent), but we did not record what was said while the document was created. We analyzed documents using three different approaches:

1. Information visualization using DocuViz [33; 34]: Using information from the data traces that make up revision histories on Google Docs, DocuViz produces a visual history chart across different time points, indicating the authors, their respective portions of writing, and the time. Figure 2 shows four such history charts, where each dark vertical bar is a slice from the revision history, showing who wrote what when, with color indicating author and height representing the amount, and position in the bar reflecting the position in the document. The shaded bars between the slices show the activity, either no changes by a continuous colored area, or changes, with a white triangle opening to the right indicating an addition, and with it opening to the left, a deletion.

In addition, this tool provides usage statistics related to collaborative revision behaviors, such as the amount of peer editing or the proportion of the final document each member wrote. Based on the trace information, we developed three variables.

- Evenness of participation: One minus the variance of proportions times 100 (to make the scores readable). Higher scores indicate higher balance in written participation. Note that since we do not have what they said, participation here refers only to how much each person *wrote*.
- Group activity: This measures the total writing and editing in a document. This is an indicator of how active a group is in writing and editing, and
- Peer-editing: This calculates the amount of editing each student did on a peer's text.

2. Computational text analysis using Coh-Metrix: In order to understand the textual characteristics of the essays, we used Coh-Metrix, a web application that analyzes linguistic and discourse features using natural language processing (NLP) techniques. Among the available Coh-Metrix indices, we selected three measures that have been most widely used and validated in writing research: lexical diversity, syntactic complexity, and textual cohesion [22; 27]. In the initial analysis, we tested multiple indices that measure the text characteristics at the lexical, syntactic, and cohesion level, but report the following indices as they showed the strongest association.

- Lexical frequency measures average word frequency for all words. It is an indicator of vocabulary breadth (i.e., size) and lexical sophistication.
- Syntactic complexity measures the number of words before the main verb.
- Textual cohesion measures an incidence score (occurrence per 1000 words) for all connectives (i.e., cohesive links between ideas and clauses). Higher scores indicate stronger connection among sentences due to frequent use of connectives.

3. Human grading of writing quality based on analytic rubrics: Finally, the essays were evaluated by the first author and a research assistant who is pursuing a graduate degree in Education and has extensive experience in teaching composition. They rated them on:

- Mechanics (i.e., proper use of spelling, punctuation, grammar),
- Content/idea (i.e., clarity of idea/thesis),
- Organization (i.e., logical structure), and
- Evidence support (i.e., appropriate and effective use of evidence).

After several rounds of practice grading sessions, the two raters graded the 45 essays on four dimensions using a 10-point scale. The inter-rater reliability using Cohen's Kappa ranged from 0.74-0.93: Mechanics (.93), Content/idea (.83), Organization (.74), and Evidence (.76).

To answer **RQ1**, we qualitatively analyzed the DocuViz visualizations to identify commonly used writing styles and noted the characteristics that identify these styles. The initial coder developed a coding category following a grounded approach [7] and invited a second coder to use these codes to categorize four visualizations. In the first cycle of coding, we generated the visualization charts of the 45 essays using DocuViz to identify distinct patterns. In the second cycle of coding, we qualitatively examined the revision histories and in-text communication to modify and confirm the categorization. The qualitative codes generated during this process was partly informed by Olson et al. [25] and included the presence of leader, member roles (e.g., writer, editor, reviewer, equal work, consultant), writing strategies (e.g., scribe, separate, synchronous), and peer editing behaviors (e.g., during or after writing). These codes were used as the main criteria for categorizing and confirming the

four visualization categories. The inter-coder reliability (Cohen's Kappa) for categorizing the visualizations was .91.

Next, the collaboration variables (e.g., participation equality, group activeness) and outcome variables (e.g., analytic quality scores, computational text analysis results) were analyzed using ANOVA to examine the differential characteristics across writing styles. To answer **RQ2**, we used multi-level regression analyses to investigate how these characteristics relate to the quality and quantity of the texts. Based on these analyses, we sought to build a holistic understanding of students' collaborative writing practices during the synchronous collaboration, and interpret the relationship between the practices and the resulting text outcome.

RESULTS

RQ 1: Collaborative Writing Styles

The following four styles of writing emerged from a bottom-up grounded analysis of DocuViz visualizations that render the collaboration history of a document: (a) *Main Writer* (Scribe in [26]) (b) *Divide and Conquer* (Horizontal Division in [26]) (c) *Cooperative Revision* (d) *Synchronous Hands-on* (see Figure 2).

In the *Main Writer* style (2a), one or two main writers dominate, writing most of the text, while the other members participate minimally. In the Figure, the person depicted in green color wrote most of the document, with the person depicted in orange color adding some and the person in red adding a tiny amount.

In the *Divide and Conquer* style (2b), they write their own parts and rarely edit each other's text. In the Figure, the writers keep to their own sections rarely crossing into something someone else wrote.

In the *Cooperative Revision* style (2c), members may divide their sections but edit each other's freely, mostly at the later stage of writing. In the Figure, Green does a lot of editing of the whole document late.

Lastly, in the *Synchronous Hands-on* style (2d), members create sentences together by simultaneously building off of each other's text. In the Figure, there is a lot of mingling of colors early, and editing each other's work happens throughout.

The *Cooperative Revision* style was most common (40%), followed by the *Main Writer* style (31%), the *Divide and Conquer* style (20%), and the *Synchronous Hands-on* style (9%). Out of the 15 groups, only six of them maintained the same style across the three documents. If it was random, only 1 team should have the same style across its three documents. So it is significantly higher than chance.

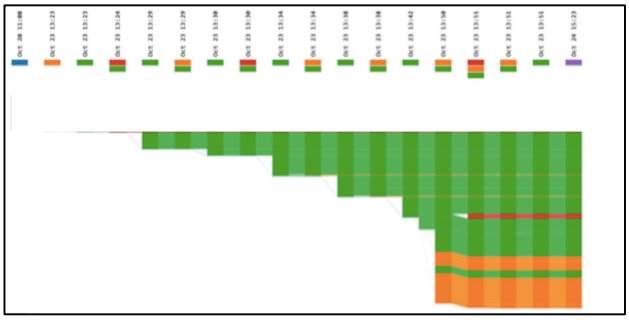

Figure 2a. An illustration of *Main Writer* style in visualization.

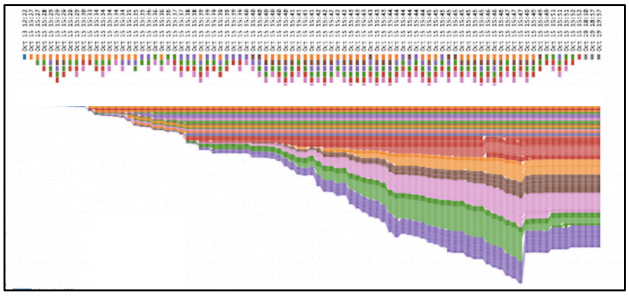

Figure 2b. An illustration of *Divide and Conquer* style.

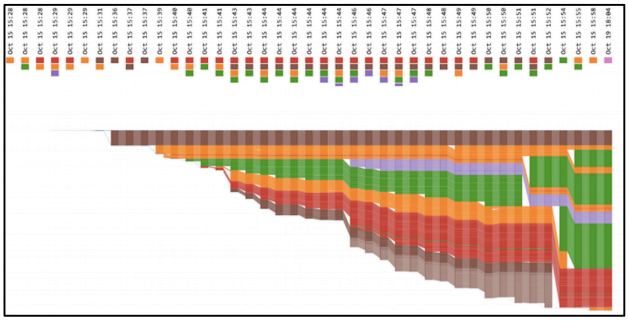

Figure 2c. An illustration of *Cooperative Revision.*

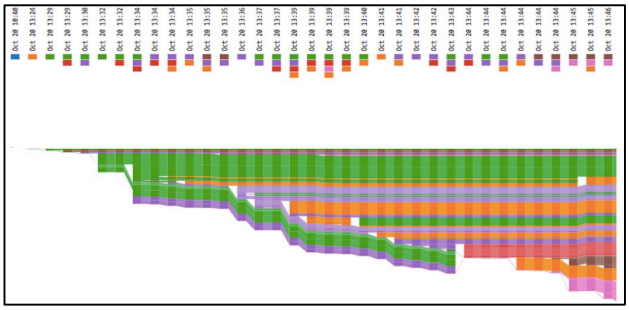

Figure 2d. An illustration of *Synchronous Hands-on.*

Differences across Styles

We used one-way ANOVA or Welch ANOVA (depending on whether the assumption of homogeneity of variance is violated) to examine each measurement to test whether there are differences in (a) collaboration characteristics, (b) text quality, and (c) quantity across the four different writing styles. When the Welch's F test revealed a statistically significant main effect, we conducted post hoc comparison using the Games-Howell post hoc procedure to determine

	Main Writer	Divide and Conquer	Cooper-ative Revision	Synchro-nous Hands-on
Collaboration characteristics				
# of Authors	4.71	5.00	4.89	6.00
Balance of Participation**	**93.59**	**98.28**	**97.69**	**99.05**
Group Activeness**	**5827.00**	**9439.67**	**9454.39**	**4719.75**
Peer Edit	753.21	600.44	1354.39	679.75
Rubric-guided quality				
Content *	**6.71**	**8.33**	**8.06**	**7.50**
Organization	7.21	7.44	7.50	7.50
Evidence**	**7.36**	**9.00**	**8.89**	**8.00**
Mechanics	8.21	8.11	7.94	9.00
Computational quality				
Textual Cohesion	86.41	95.54	89.53	104.75
Syntactic Complexity	5.63	5.03	4.97	4.85
Lexical Frequency	2.88	2.91	2.93	2.83
Text quantity				
# of Paragraphs **	**2.86**	**6.11**	**3.94**	**1.5**
Paragraph Length	6.42	6.21	8.01	9.88
# of Sentences **	**17.14**	**36.89**	**28.61**	**13.5**
# of Words **	**347.07**	**751.33**	**576.39**	**257.5**

Table 2. ANOVA results: differences in means across writing styles (bold indicates significant differences).

which pairs of the four style means differed significantly. Table 2 summarizes the results.

The first two are minor, mere confirmations of initial categorizations of writing styles based on DocuViz visualizations. The more interesting ones follow.

Differences in collaboration characteristics

***a. Even participation*:** The results indicated that documents written with *Main Writer* style were significantly less balanced than documents written with the other three styles, confirming the dominance of one or two people. Additionally, documents written in the style of *Cooperative Revision* were significantly less balanced than the documents with *Synchronous Hands-on* style.

b. Editing behavior: The four styles significantly differed in group activeness (i.e., total writing and editing). Documents written with *Cooperative Revision* and *Divide and Conquer* style had higher levels of activeness than documents written with *Main Writer* or *Synchronous Hands-on* style.

Differences in text quality

a. Rubric-guided quality scores: There were statistically significant differences on the *Content* and *Evidence* scores in different styles of writing. Content was rated lower in documents with a *Main Writer* style than in documents with *Divide and Conquer* style or with *Cooperative Revision* style. Evidence was also rated lower in documents with a *Main Writer* style than *Divide and Conquer* and *Cooperative Revision*. Although *Divide and Conquer* appears as the highest in both of these, it was not significantly higher than *Cooperative Revision* or *Synchronous Hands-on.*

b. Computational measures from Coh-Metrix: The styles were not significantly different on any of the Coh-Metrix measures.

Differences in text quantity

We found significant difference on the measurements of *Number of Paragraphs*, *Number of Sentences*, and *Number of Words*, but not on the *Paragraph Length*. Post hoc comparisons indicated that the *Number of Paragraphs* was higher in documents written with *Divide and Conquer* style than in documents with other styles. Documents written in *Cooperative Revision* style had more paragraphs than those written in *Synchronous Hands-on* style. The styles of *Main Writer* and *Synchronous Hands-on* produced a smaller *Number of Words and Sentences* than the styles of *Divide and Conquer* or *Cooperative Revision*.

RQ 2: What Relates to Document Quality and Quantity?

We used multi-level regression analyses to examine how collaboration-related characteristics (e.g., even participation, group activeness) relate to text outcome measures that were extracted from both analytic grading and computational text analysis. Our analyses revealed that collaboration characteristics relate to both text quality and quantity (Figures 3 and 4). We then examined whether different measures of text quantity relate to quality (Figure 5). And, finally, we measured the relationship between the two kinds of quality measures (Figure 6).

Impacts on Document Quality

Effects of number of authors*:* As shown in Figure 3, more authors led to better Mechanics (β=0.52, p<.01), which implies that the group may use additional help and attention available from having more members to polish the mechanics of the paper. This is the only relationship that involves either the number of authors or Mechanics.

Effects of even participation: Balanced participation predicted higher scores in Content (β=0.27, p<.01), Evidence (β=0.26, p<.01), but not Organization and Mechanics. Balance is negatively related to Syntactic Complexity (β=-0.31, p<0.01). We found no significant relationships in Lexical Frequency or Textual Cohesion.

Effects of editing behavior: Group activity (i.e., the amount of writing and editing in a document) predicted

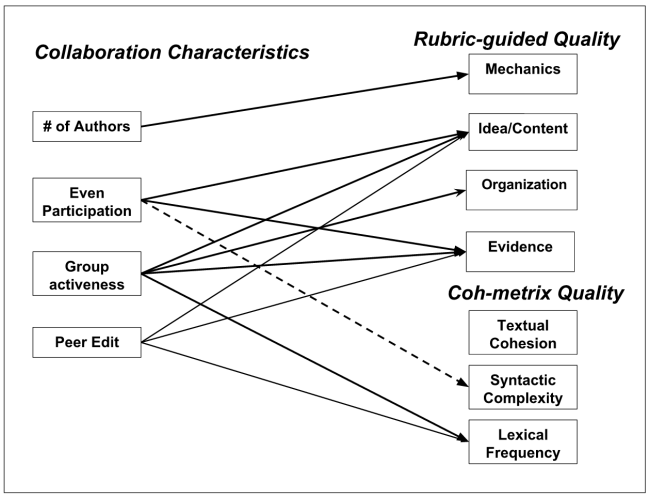

Figure 3. The relationship between collaboration characteristics and both kinds of quality. Dotted lines indicate negative relationships.

higher scores in all dimensions except Mechanics: Content (β=0.00021, p<.01), Evidence (β=0.00021, p<.01), Organization (*β*=.00018, *p*<.01). The amount of peer editing predicted higher scores in Content (*β*=.00043, *p*<.05), Evidence (*β*=.00039, *p*<.05), but not in Organization or Mechanics. Among the computational linguistic traits, only the Lexical Frequency was related to editing behaviors. Both group activity (*β*=.000008, *p*<.05), and peer editing (*β*=.00003, *p*<.05) predicted stronger Lexical Frequency. The more people write and edit the document as a group, and edit their peers' text, the more diverse vocabulary is used in the group document.

Impacts on Quantity

Effects of number of authors: As shown in Figure 4, having more authors did not relate to document length. This was interesting as we hypothesized that groups with more members might produce longer documents. This implies that the length of documents relies more on the even level of participation rather than merely the number of authors.

Effects of balance: Even participation predicted more sentences (*β*=1.55, *p*<.01), and words (*β*=25.36, *p*<.05) in a text, but not paragraph length or number of paragraphs in the text. When every participant is contributing to the content (balanced participation), they produce more words and sentences, but not necessarily more paragraphs, suggesting that they write longer sentences.

Effects of editing: When members are more active in writing and editing (i.e., group activeness), the more words are in the documents (*β*=.038, *p*<.01) and sentences (*β*=.0021, *p*<.01) level. Group activeness predicted more paragraphs (*β*=.00016, *p*<.05), but not the longer paragraphs. Interestingly, peer editing did not predict any of the length measures. This may imply that peer editing may involve some minor changes, but not a significant addition of content to the text not written by oneself, potentially due to psychological ownership of the text.

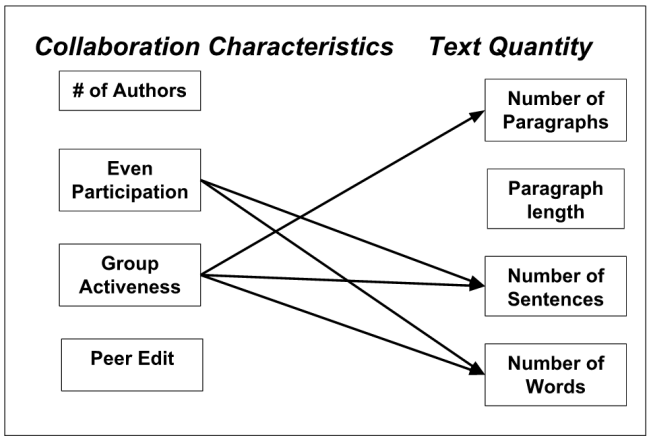

Figure 4. Relationship between collaboration characteristics and document quality.

Relationship Between Quantity and Quality

In addition, we examined whether text quantity relates to quality (Figure 5). Our analyses revealed that both number of words and sentences in a text predicted higher scores in Content (*β*=.0029, *p*<.01; *β*=.053, *p*<.01, respectively), Evidence (*β*=.0032, *p*<.01; *β*=.062, *p*<.01), Organization (*β*=.0020, *p*<.01; *β*=.041, *p*<.01), and Lexical Frequency (*β*=.00012, *p*<.05; *β*=.0019, *p*<.05). Number of sentences was negatively related to Syntactic Complexity (*β*=-0.051, *p*<.05). Number of paragraphs used was generally a stronger predictor for quality than paragraph length. It predicted higher scores in Content (*β*=0.27, *p*<.05), Evidence (*β*=0.26, *p*<.01), and Lexical Frequency (*β*=0.019, *p*<.01), whereas paragraph length related only to Evidence (*β*=0.26, *p*<.01). None of the quantity measures were related to the Rubric-guided quality measure of Mechanics, or Coh-Metrix's Syntactic Complexity or Textual Cohesion.

Relationship Between Two Kinds of Quality

Figure 6 shows the relationship between the two kinds of quality, that from human judgment following a rubric and that from Coh-Metrix algorithms. Interestingly, all of the

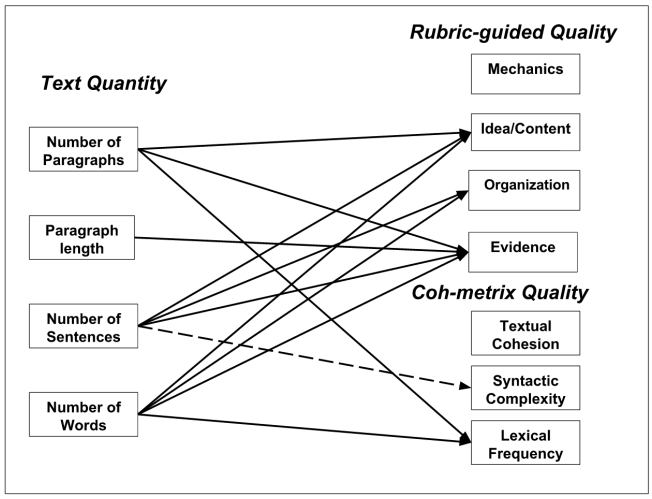

Figure 5. The relationship between quantities and both kinds of quality.

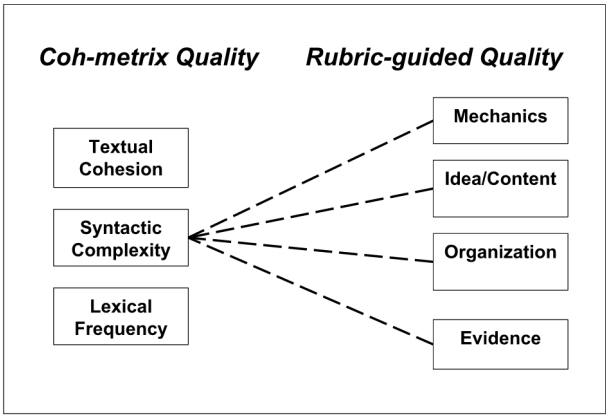

Figure 6. The relationship between the two measures of quality.

rubric-guided measures were significantly correlated with Syntactic Complexity, and all negative (Mechanics: β=-0.603, p<.01, Idea/Content: β=-0.575, p<.01, Organization: β=-0.522, p<.01, Evidence: β=-0.687, p<.01). The simpler the structure, the better the Mechanics, Ideas, Organization and Content are rated.

DISCUSSION

In this study, we examined the undergraduate students' synchronous collaboration practices in classroom exercises, and found notable impacts of collaboration characteristics (i.e., participation equality, group activeness, amount of peer edit) on writing style, writing quality, and quantity. Among the multiple findings, we want to highlight that:

- *Divide and Conquer* style tended to produce better quality text, particularly in content and evidence, whereas *Main Writer* style had the lowest scores in those areas.
- Balanced participation and active editing behaviors predicted better writing quality (e.g., content, evidence, lexical frequency) and quantity.

These results have several significant implications for both undergraduate instruction and system design.

Educational Implications

First, we found four different styles of synchronous writing: *Main Writer, Divide and Conquer, Cooperative Revision, and Synchronous Hands-on*. It was interesting to see that some groups explored new ways of writing together by building off each other in writing sentences synchronously (i.e., *Synchronous Hands-on* style). However, most of the groups wrote in familiar ways of working together (*Main Writer, Divide and Conquer)*. While *Divide and Conquer* or *Cooperative Revision* can be comparable to traditional styles such as parallel writing or peer review practices, the synchronous writing and editing enabled the students to produce significantly longer documents than the other two styles, which may indicate a sign of task efficiency. Text quantity can also be another proxy indicator of how much effort a team puts into writing, or how efficiently a team worked to produce a text.

What is more intriguing is the link between writing style and text quality. In our study, *Divide and Conquer* style was third most common but exhibited the highest participation equality and high though not the highest group activity. It also had higher quality in terms of content and evidence use. Interestingly, *Main Writer* style produced lowest scores in content and evidence, which indicates that writing that involves a dominant writer may not lead to higher writing quality. Given that balanced participation, as evident in the *Synchronous Hands-on, Cooperative-Revision* or *Divide and Conquer* styles, is associated with strong content and evidence use, instructors may explore teaching or task design strategies to encourage both balanced and active participation. One can imagine any activity that is done simultaneously, like brainstorming by silently typing, could encourage more even participation and more work.

We are not ready to recommend that people be encouraged or required to engage in these activities, because we do not yet know the causal connection. It may be, for example that the more capable students naturally navigate towards these behaviors and styles. Similarly, perhaps when one capable student is in a group with those less capable, they will dominate, but not do as well as a group of all highly capable students. Now that we know the relationships exist, future research can investigate the causal connection. We could assess students' GPAs or another academic performance evaluation metric.

In addition, more research is needed to investigate what factors may contribute to different participation or collaboration patterns both at the individual or group level. For example, we need to learn more about how socio-emotional factors (e.g., writing anxiety, writing efficacy, group cohesion) or individual characteristics (e.g., familiarity with technology, experience with collaborative writing, trust in each other) may impact group dynamics, students' writing behaviors or styles in synchronous collaboration mode.

Next, the findings on document quality suggest that synchronous, collocated writing may enhance the content and evidence of text, but not in organization and mechanics. Our results indicate that balanced pooling of ideas from multiple authors in synchronous contexts may only strengthen the content, yet when careful attention is needed to polish the organizational structure of text, that work should be done solo. That is, the presence of peer-readers in synchronous collaborative context does not necessarily enhance the organization, which contradicts findings from previous studies on asynchronous feedback practices [20; 12] including those on wikis; careful efforts are necessary to tie together different ideas pooled from members. Given that most of the previous findings examined asynchronous collaboration, it is also possible that writers may tend to become careless or over-dependent on peer support in synchronous mode of collaboration, or have little ownership of the whole. Therefore they may not pay sufficient

attention to the organizational or mechanical aspect during the final polishing/revision stage.

Task type may also have impacted the collaborative outcomes in our study. The time given for the task (50 min) may have been limited and thus may not have allowed members sufficient time to work on revision for improving organization. Our results highlight the importance of considering task design, particularly time assignment, and the need to explicitly integrate a revision stage in the group writing tasks [2]. Tasks that combine synchronous, in-class collaboration with asynchronous, independent activities might be helpful for facilitating organization and review/revision stages. This is especially so considering that collaborative writing should not be taken as merely an efficient strategy for task completion without little collaboration (e.g., *Divide and Conquer*), particularly in instructional settings, where the goal is not productivity, but learning the educational content or writing skills. Rather, to enhance the educational value of peer feedback and edits, instructors and system designers should support creating the conditions in which they can be effective, for example, by guiding them through collaboration stages or raising their awareness of collaborative behavior and writing process. In addition, future research may also investigate whether specific task structure (e.g., pre-writing activities, planning, revision stage), genres (e.g., narrative, argumentative, informative) or group structure (e.g., ability grouping) may facilitate certain styles of collaboration or promote equal participation (e.g., whether the absence of planning promotes the main writer style), and what subsequent effects it has on group writing outcomes.

Design Implications

Our findings also suggest two design implications for collaborative writing systems. They could all be part of an Integrated Writing Environment (IWE), similar to the Integrated Development Environments (IDE) for programming teams like Eclipse: a visualization of participation and a meter of writing quality. The two design implications correspond to different aspects of our findings.

First, we propose to provide a visualization to reveal each individual co-author's participation in real time. Our result shows that *even participation* is associated with several dimensions of higher quality of the final document. If future research confirms a causal relationship, we could encourage even participation with a visual display of how even the participation is up to this moment. The design could be a dashboard or a summary table with team members' names and their contributed characters, similar to DocuViz's report, but presented in real time, not by running DocuViz. By visualizing the activity level of each group member, we expect to promote higher level of participation and engagement from all the group members and to reduce the number of free-riders.

Second, Coh-Metrix can be used to assess the quality of the writing (e.g., by using Text Easability Assessor that compares the Coh-Metrix scales in the given sample to large corpora means; http://tea.cohmetrix.com/). We found that different features of editing behavior are related to these quality measures: Both Group Activeness and Peer Editing are related to Lexical Frequency. Therefore, we propose that the IWE could calculate and present some dimensions of quality when requested, along with suggestions about how to improve the writing if they are out of bounds on that characteristic. Previous research has also suggested that providing early feedback to writers can trigger their reflection and revision behaviors, which can lead to higher quality [31]. We expect the design feature of calculating and presenting the quality aspects in an IWE can trigger co-authors' discussion and reflection, and further improve the quality of the document.

CONCLUSION

In this study, we revealed how undergraduate students simultaneously write and collaborate using Google Docs in a classroom, how their practices affect the text quality and quantity, and what implications the results may have for both system design and instruction. Our findings highlight the importance of balanced participation and group activity in writing a better quality text, particularly in the areas of content and evidence, and also in working efficiently to produce longer documents. By incorporating multiple methods including information visualization, computational text analysis, and rubric-guided quality assessment, we provided a finer grained analysis of the link among synchronous writing styles, quality, and quantity.

As we suggest, system designs that help raise students' awareness on both collaboration (e.g., participation levels) and aspects of quality have the potential of supporting instruction and helping to maximize the educational benefits of synchronous collaboration. We hope that our work triggers an intriguing line of future work to help people write better together by providing appropriate support for them to harness the power of simultaneous writing.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their thoughtful and constructive comments. This work is supported by a Google Focused Faculty Research Award and a grant from the National Science Foundation (ACI-1322304).

REFERENCES

1. Recep S. Arslan, and Aysel Şahin-Kızıl. 2010. How can the use of blog software facilitate the writing process of English language learners? *Computer Assisted Language Learning*, 23(3), 183-197.
2. Zeliha Aydin, and Senem Yildiz. 2014. Using wikis to promote collaborative EFL writing. *Language, Learning & Technology* 18, 1: 160.
3. Ronald M. Baecker, Nastos Dimitrios, Ilona R. Posner, and Kelly L. Mawby. 1993. The user-centered iterative design of collaborative writing software. In

Proceedings of the INTERACT' 93 and CHI' 93, 399-405.

4. Jeremy Birnholtz and Steven Ibara. 2012. Tracking changes in collaborative writing: edits, visibility and group maintenance. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, 809-818.
5. Jeremy Birnholtz, Stephanie Steinhardt, and Antonella Pavese. 2013. Write here, write now!: an experimental study of group maintenance in collaborative writing. In *Proceedings of the CHI'13*, 961-970.
6. Samuel Kai-Wai Chu, and David M. Kennedy. 2011. Using online collaborative tools for groups to co-construct knowledge. *Online Information Review* 35, 4: 581-597.
7. Juliet Corbin, and Anselm Strauss. 1990. Grounded theory method: Procedures, canons, and evaluative criteria. *Qualitative Sociology* 13: 3-21.
8. Scott A. Crossley, Jerry Greenfield, and Danielle S. McNamara. 2008. Assessing text readability using cognitively based indices. *TESOL Quarterly* 42, 3: 475-493.
9. Richard Donato. 1994. Collective scaffolding in second language learning. *Vygotskian approaches to second language research*, 33-45.
10. Paul Dourish, and Victoria Bellotti. 1992. Awareness and coordination in shared workspaces. In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, 107-114.
11. Lisa S. Ede, and Andrea A. Lunsford. 1990. Singular texts/plural authors: Perspectives on collaborative writing. SIU Press.
12. Idoia Elola and Ana Oskoz. 2010. Collaborative writing: fostering foreign language and writing conventions development. *Language Learning & Technology* 14, 3: 51-71.
13. Arthur C. Graesser, Danielle S. McNamara, Max M. Louwerse, and Zhiqiang Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers* 36, 2: 193-202.
14. Henry Jenkins. 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. Cambridge, MA: The MIT Press.
15. Robert Johansen. 1989. User approaches to computer-supported teams. In *Technological support for work group collaboration*, Margrethe H. Olson (Ed.). L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1-31.
16. Greg Kessler, and Dawn Bikowski. 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.
17. Hee-Cheol Ezra Kim, and Kerstin Severinson Eklundh. 2001. Reviewing practices in collaborative writing. *Computer Supported Cooperative Work*, 10: 247-259.
18. Muhammad M. Latif. 2012. What do we mean by writing fluency and how can it be validly measured? *Applied linguistics*, 1-8.
19. Paul Benjamin Lowry, Aaron Curtis, and Michelle René Lowry. 2004. Building a taxonomy and nomenclature of collaborative writing to improve interdisciplinary research and practice. *Journal of Business Communication* 41, 1: 66-99.
20. Barley Mak and David Coniam. 2008. Using wikis to enhance and develop writing skills among secondary school students in Hong Kong. *System*, 36, 3: 437-455.
21. Lola J. McGuffin, and Gary M. Olson. 1992. *ShrEdit: A Shared Electronic Work Space*. University of Michigan, Cognitive Science and Machine Intelligence Laboratory.
22. Danielle S. McNamara, Yasuhiro Ozuru, Arthur C. Graesser, and Max Louwerse. 2006. Validating coh-metrix. In *Proceedings of the 28th annual conference of the cognitive science society*, Mahwah, NJ: Erlbaum, 573-578.
23. Alex Mitchell, Ilona Posner, and Ronald Baecker. 1995. Learning to write together using groupware. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, 288-295.
24. Judith S. Olson, Gary M. Olson, Marianne Storrøsten, and Mark Carter. 1993. Groupwork close up: A comparison of the group design process with and without a simple group editor. *ACM Transactions on Information Systems (TOIS)* 11, 4: 321-348.
25. Judith S. Olson, Dakuo Wang, Jingwen Zhang, and Gary M. Olson. 2016. How people write together now: Beginning the investigation with advanced undergraduates in a project course. Under review.
26. Ilona R. Posner, and Ronald M. Baecker. 1992. How people write together [groupware]. In *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on* 4: 127-138.
27. Wenjuan Qin, and Paola Uccelli. 2016. Same language, different functions: A cross-genre analysis of Chinese EFL learners' writing performance. *Journal of Second Language Writing*, *33*, pp.3-17.
28. Roy Rada. 1996. *Groupware and Authoring*. Academic Press.
29. Yu-Chih Sun, and Yu-jung Chang. 2012. Blogging to learn- Becoming EFL academic writers through collaborative dialogues. *Language Learning & Technology* 16, 1: 43-61.

30. Yunting Sun, Diane Lambert, Makoto Uchida, and Nicolas Remy. 2014. Collaboration in the cloud at Google. In *Proceedings of the 2014 ACM conference on Web science*, 239-240.
31. Neomy Storch. 2005. Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing* 14: 153-173.
32. Carola Strobl. 2014. Affordances of web 2.0 technologies for collaborative advanced writing in a foreign language. *Calico Journal* 1: 1-18.
33. Dakuo Wang, Judith S. Olson, Jingwen Zhang, Trung Nguyen, and Gary M. Olson. 2015. DocuViz: Visualizing Collaborative Writing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1865-1874). ACM.
34. Dakuo Wang. 2016. Exploring and Supporting Today's Collaborative Writing. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 255-259). ACM.
35. Saovapa Wichadee. 2011. Using wikis to develop summary writing abilities of students in an EFL class. *Journal of College Teaching & Learning* 7, 12: 5-10.
36. Luxin Yang. 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.
37. Hui-Chin Yeh. 2014. Exploring how collaborative dialogues facilitate synchronous collaborative writing. *Language Learning & Technology* 18, 1: 23–37.