Affordances of Web 2.0 Technologies for Collaborative Advanced Writing in a Foreign Language

CAROLA STROBL
Ghent University, Belgium

ABSTRACT
Can online collaboration yield a positive effect on academic writing in a foreign language? If so, what exactly is the value, compared to individual writing, and (how) does it translate to better output? These are the central questions addressed in this paper. Second language (L2) writing research has long highlighted the benefits of collaboration in terms of both L2 learning and text quality. Most recently, the positive effect of co-ownership and peer feedback on process and product has been emphasized in studies on computer-supported collaborative writing. What has remained underexplored is the impact of Web 2.0 technologies on advanced L2 writing. The present paper bridges this gap through an empirical study combining Web 2.0 technologies with an academic writing task. Collaborative and individual writing processes and products are compared by applying a mixed-methods approach. Results shed new light on claims made in previous studies that collaboration leads to higher text accuracy. No statistically relevant difference was found between the individual and collaborative syntheses in terms of complexity, accuracy, and fluency. However, collaborative texts score significantly higher on appropriate content selection and organization. Analysis of the process data shows that this is due to in-depth discussions during the planning phase.

KEYWORDS
Computer-supported Collaborative Writing (CSCW), Advanced Foreign Language Writing, Revision Process, Text Quality, Peer Feedback

INTRODUCTION
The task under investigation in this study, writing a synthesis from several written sources in a foreign language, is highly complex because it requires a multiple focus on input processing, content selection, and organization, along with production. Can online peer collaboration help students to master it? Web 2.0 technologies offer great opportunities for the collaborative co-construction of knowledge. Wikis and online editors facilitate the co-construction of a joint artifact, in this case, a collaboratively written text. Due to increasing levels of sophistication and usability, these tools have received growing interest from writing instructors and researchers alike. Is it possible, though, to find empirical evidence for the supposed benefit of the online collaboration process in the product, in this case, a synthesis from written sources? Evidence-based research on this question has remained scarce to date (Wang & Vásquez, 2012). This paper answers the need for an empirical study of the benefits of Web 2.0 tools for advanced L2 text production.

The research idea for this study is driven by two complementary lines of scholarship: L2 writing research, and computer supported collaborative writing (CSCW). L2 writing research has long highlighted the benefits of collaboration in terms of both L2 learning and text quality (Arnold, Ducate, & Kost, 2009; De la Colina & García Mayo, 2007; Elola & Oskoz, 2010; Fernández Dobao, 2012; Kessler, Bikowski, & Boggs, 2012; Kost, 2011; Storch, 2005; Wigglesworth & Storch, 2009). Where text quality is concerned, texts resulting from
pair work have been found to achieve a higher level of accuracy than those written by individuals (Storch, 2005; Wigglesworth & Storch, 2009). Fernández Dobao (2012) even found a significant difference in accuracy dependent on the number of participants involved in the collaborative process (pairs vs. groups of three). She attributed this difference to the higher amount of correctly resolved language-related episodes in the groups of three.

Besides product enhancement, a possible connection between the pooling of linguistic resources in collaborative dialogues and L2 learning has been explored (for a comprehensive review see Swain, Brooks, & Tocalli-Beller, 2002). Studies in CSCW have emphasized the positive effect of co-ownership and electronic peer feedback regarding reader-orientation (Blin & Appel, 2011; Kessler, 2009; Lee, 2010). Departing from an Activity Theoretical framework, the emerging text (a "digital material-semiotic artifact") is interpreted simultaneously as an object and as a tool, and therefore, research interest has shifted from the product toward the writing process: "digital material-semiotic artifacts ... provide data not only on language use and students’ interactions, but also on the evolving social and communicative structure of the activity, or the hidden curriculum, as it is negotiated and possibly re-constructed by the participants as the activity unfolds" (Blin & Appel, 2011, p. 477). Specific attention has been given to the role of feedback during the writing and reviewing process, particularly formative feedback, which can either be given by the teacher or by peers. In the former case, the teacher "shares the responsibility to assess with the students, but still supervises the revision process" (Alvarez, Espasa, & Guasch, 2012, p. 388), whereas in the latter, this responsibility is shared among peers. In addition, the process can be supported by the teacher through a script that provides instructions on how to interact, promoting a fruitfully structured interaction (Rummel, Spada, & Hauser, 2009).

Peer collaboration in the reviewing process has been found to focus mainly on surface-level, meaning preserving changes (Paulus, 1999), also called LOCs (lower-order concerns). What has remained underexplored, however, is the impact of Web 2.0 technologies on advanced L2 writing (Thorne & Reinhardt, 2008) where HOCs (higher-order concerns) like content selection and organization play an important role. The present paper aims to bridge this gap through an empirical study. In this study, the use of Web 2.0 technologies throughout the writing process is combined with a task, synthesis writing, which belongs to the field of advanced foreign language proficiency and imposes a high cognitive load on the learners. The research questions and sub-questions for the present study are:

(1) What impact does online collaboration have on the final text?
   • Do the collaborative syntheses differ from individually written ones with regard to complexity, accuracy and fluency (CAF) and/or content and coherence?
   • Do collaboratively produced texts (mainly) achieve higher accuracy levels?

(2) What impact does online collaboration have on the writing process?
   • How does the collaborative activity influence the writing and reviewing patterns, compared to individual writing?
   • Is it true that peer-induced revision mainly focuses on surface-level, meaning-preserving changes, that is, LOCs?

METHOD

In order to explore the difference between collaborative and individual online writing, an empirical study was set up using a university virtual learning environment and Google Docs as tools. The research was carried out in an intact class (n=48) of university students of German L2. All of them were Dutch native speakers and had an advanced proficiency level of German writing (CEF-level B1-B2). Prior to the tasks under investigation, all students received an introduction to synthesis writing and to peer reviewing. The instructor modeled the revision of several authentic students’ syntheses from a different class in a guided classroom activity.

For the purposes of the actual study, the class was divided into two groups. All participants were required to write two syntheses in crossed conditions, one individually,
and one in a randomly selected collaborative group of three, one from aural, and one from written input (see Table 1 for an overview). As the two synthesis tasks differed in terms of input type and outcome, only one of them, synthesis based on written sources, will be reported on in this paper.

**TABLE 1**

Overall Study Design: Two Tasks under Crossed Conditions

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Introduction to synthesis writing and peer reviewing</td>
<td>Pre-hoc questionnaires on learning style and attitudes</td>
</tr>
<tr>
<td>Week 2</td>
<td>Synthesis based on written sources*</td>
<td>Individual writing</td>
</tr>
<tr>
<td></td>
<td>Collaborative writing</td>
<td>Individual writing</td>
</tr>
<tr>
<td>Week 3</td>
<td>Summary based on aural source</td>
<td>Collaborative writing</td>
</tr>
<tr>
<td>Week 4*</td>
<td>Post-hoc questionnaire on attitudes</td>
<td></td>
</tr>
</tbody>
</table>

* Results reported in this paper

In this task that prepares students for academic writing, they were asked to integrate and synthesize information from multiple sources on a complex topic into a coherent and informative text of a prescribed length. This means that they needed to thoroughly select the content of the input material for their own text. The source material for this task, however, was not of an academic nature. It consisted of three different pieces of textual information related to the planned construction of a new station in Stuttgart (*Stuttgart21*) which had been the subject of a considerable public debate in Germany: 1) a list of bullet points providing background on the timelines of the project, its ecological and economic impact, and the acts of civil protest it provoked, 2) the declaration of the Association for the German Language (GfdS) about the word of the year 2010, *Wutbürger*², and its relatedness to civil protest gaining momentum in Germany, and finally 3) a selection of quotes, representing a variety of viewpoints on the project.

The planning phase was organized based on constructivist learning principles, according to which knowledge is constructed in social or cognitive interaction. In the collaborative groups, social interaction was achieved among the peers through a script with stepwise instructions and an online forum. The individual writers needed a different, cognitive stimulus to interact with and to challenge their ideas. For this purpose, the individual writers followed an online exercise module with sample solutions: After submitting an answer, they were immediately prompted to compare it with the respective sample solution. In this way, they were given the opportunity to profit from the sample solution in the following step of the planning phase.

The collaborative groups were supported throughout the planning and writing process by a script with stepwise instructions helping them organize their collaboration. The effectiveness of scripts for collaboration in computer-mediated settings has been demonstrated by CSCL research. As Rummel et al. (2009) state: “By enforcing specific kinds of activities among the collaborators, scripts are expected to prompt cognitive and social processes by participants that might otherwise not occur” (p. 73). According to Dillenbourg and Hong’s (2008) classification, the script used in this study can be considered a micro-script in that it prescribes the collaborators’ actions down to the level of the dialogue.
The exploration of the source material was conceptualized as a jigsaw exercise in the collaborative groups. Each member was asked to explore one source and explain it to the others in an online forum. The individual writers were guided in this phase by reflective questions in the online exercise module. Next, both individual and collaborative writers received online guidance to plan their own texts, leading them to build a "text skeleton," that is, an outline interspersed with cohesive ties. After submitting their own skeleton proposal, the individual writers received two sample skeletons. The collaborative groups did not receive sample solutions.

The planning phase of approximately 45 minutes took place in class but was completely online. The writing and reviewing phases were carried out by all participants outside the classroom in Google Docs, an online document-sharing and versioning tool, with a one-week deadline for the delivery of the final text. Again, the collaborative groups were supported by a micro-script guiding the interaction along a timeline. The individual writers received general guidelines for the writing and review stages.

To measure the intervention outcome in terms of product and process, a mixed-methods approach was adopted. For product analysis, the classical CAF triad was measured in a semi-automated way. Complexity was computed using an algorithm by Schulze, Wood, & Pokorny (unpublished manuscript) that combines four textual surface measures into one numerical value for so-called "balanced complexity" (cb): (1) lexical variation (Giraud´s type-token ratio), (2) lexical sophistication (mean word length), (3) syntactic variation (unique bigram ratio), and (4) syntactic sophistication (mean period unit length). Accuracy was calculated as errors per word ratio (E/W), and fluency as total number of words (# W). In addition, the final texts were evaluated regarding readability by assigning a joint holistic score on a scale of 1 to 5 for coherence and cohesion. To measure content selection, all so-called "missing propositions," that is, main propositions from the source texts that had not been included in the final text were counted. Independent samples t-tests on comparison of means were conducted with a view to detecting statistically significant differences between the texts produced individually and collaboratively. To control for the small data set size, non-parametric tests (Mann-Whitney U test) were also run.

Insight into the writing process was gained through both qualitative and quantitative instruments. For the qualitative analysis, the Google Docs revision history of all seven collaborative texts and seven individual texts was traced. Furthermore, the online communication in the collaborative groups was observed analyzing the forum posts (used for planning) and the comments made in Google Docs (used for writing and revision). In addition to this qualitative analysis, the collaboration activity level in the groups during writing and revision was measured by counting the discussions consisting of more than two turns and all comments in Google Docs (both language and non-language related contributions). For a taxonomy of the collaboration activities, the comments were assigned to twelve topic-based categories. This taxonomy draws upon previously established rubrics in order to describe revision behavior (Arnold et al., 2009; Elola & Oskoz, 2010), adapting them to the data and focus of the study. According to the topics that were found to play a predominant role in the peer contributions, some LOC-items were bundled into one category (e.g., punctuation and spelling), and the categories "cohesive ties choice" and "semantics" were added. The revision type (addition, change, or deletion) was neglected because the present study focuses on revision content (HOC vs. LOC).

As groups were randomly assembled, a possible influence of the group constellation in terms of proficiency level on the collaboration activity level and/or the final outcome of the intervention was investigated using a non-parametric correlation test (Spearman´s Rho). The mean of the individual group members’ proficiency scores represents the group proficiency score. The individual scores were calculated as a combination of several previous study results related to language proficiency. To measure the final outcome, a holistic score was assigned to all final drafts, both the collaborative and the individual ones.

Furthermore, the answers of the students in the evaluative post-hoc survey were analyzed both quantitatively and qualitatively and provided additional insight into their experiences with, and attitudes toward, the two writing conditions.
RESULTS

Product Analysis

Quantitative results regarding text quality

In this section, the results of the quantitative analyses regarding CAF, on the one hand, and content selection and presentation, on the other hand, will be presented. An overview of all numeric results can be found in Table 2.

<table>
<thead>
<tr>
<th></th>
<th>Individual texts (n=21)</th>
<th>Collaborative texts (n=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Min</td>
</tr>
<tr>
<td>cb (complexity)</td>
<td>39.23</td>
<td>34.93</td>
</tr>
<tr>
<td>E/W (accuracy)</td>
<td>.08</td>
<td>.03</td>
</tr>
<tr>
<td># W (fluency)</td>
<td>327</td>
<td>260</td>
</tr>
<tr>
<td>Missing propositions</td>
<td>2.64</td>
<td>1</td>
</tr>
<tr>
<td>Cohesion + coherence</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Holistic score (/20)</td>
<td>12.3</td>
<td>8</td>
</tr>
</tbody>
</table>

*Statistically significant difference at the .05-level between individual and collaborative texts (t=-2.652, df=24.76, p=.014)

** Statistically significant difference at the .05-level between individual and collaborative texts (t=2.080; df=26, p=.045)

The descriptive statistics suggest that the collaborative texts (n=7) achieve higher levels of accuracy, fluency, and content selection than the individual texts (n=21). On average, collaboratively written texts contain fewer errors, more words, and there are fewer main propositions missing. The E/W ratio of the individual texts averaged .08, ranging from .03 to .13. In the collaborative texts, the range is smaller (from .02 to .1), and the mean is lower (.06). Regarding text length, the difference in range is even larger (203 versus 26 words).

For all text quality categories, a t-test on comparison of means was performed in order to detect a possible statistical significance which was set at a p-value <.05. In order to account for the small overall sample size and the difference in size between the two independent samples, a non-parametric Mann-Whitney U test was also run which confirmed the results of the t-test. The difference in fluency and content selection is statistically significant, but this is not the case with accuracy (t=1.201, df=26, p=.241). Moreover, individual and collaborative texts neither differ in cohesion and coherence (t=.311, df=26, p=.759), nor in balanced complexity (t=.113, df=26, p=.911).

Process Analysis

Qualitative analysis of the planning phase: Content selection

The results presented in this section summarize the qualitative analysis of two data sources: a comparison between the text skeletons and the final text drafts, and the planning
conversations in the forums. In order to detect the influence of social and cognitive stimuli on the planning process, the text skeletons were compared to the final text drafts in terms of proposition selection and order. The following example illustrates the difference between the two writing conditions (individual and collaborative) in terms of content selection. One of the source materials to be integrated into the synthesis was the declaration of the GfdS concerning the selection of *Wutbürger* as word of the year. This macro-proposition is not easy to integrate into the rest of the source material, and was suggested as an introductory or conclusive item in the sample solutions to the individual writers. Nevertheless, 70% of them chose not to integrate the proposition into their final texts and simply dropped it. This is in stark contrast to the collaborative groups, all of whom integrated the proposition into their texts. The following excerpt of a forum discussion reflects the kind of difficulties they encountered while trying to fit the proposition into their text skeletons.

Excerpt 1: Forum discussion (English translation followed by the German original, italicized):

S1: I would not mention *Stuttgart 21* as example for „*Wutbürger*“, but just as a stand-alone proposition. After all, the text does not depart from the concept of „*Wutbürger*“, but from S21. The „*Wutbürger*“-concept and its implications could then be used as an introduction.

→S2: You are right. In my skeleton proposal it looks like word of the year was the main topic instead of *Stuttgart 21*.


→S2: *Damit haben sie Recht. Wie ich es in diesem Textskelett vorschlage, sieht es ja so aus, als ob der Text um das Wort des Jahres gehen wird, statt um Stuttgart 21.*

**Qualitative analysis of the writing process**

As the writing and reviewing phases are very much intertwined in a recursive writing process, they will be treated simultaneously in this section. This approach seems particularly suited for the description of the asynchronous collaborative writing process where the two phases are hardly distinguishable: while one group member still engages in writing a first draft of a text passage, another member may be reviewing already written text of the same document simultaneously. This means that often, emerging text and already written and/or revised text are synchronously present in the same document.

At this point, a caveat is in order. Although the students were instructed to write all texts in Google Docs, some of them clearly preferred to use their usual text editor, and subsequently cut and pasted their individual contributions, or even their complete final drafts, into the Google Docs editor. This is true for about 80% of the individual texts, where only a final edit, or no action at all, can be tracked via the Google Docs history function. This being the case, no quantitative conclusions regarding revision cycles and activities can be drawn. Instead, a qualitative analysis of the text geneses that could be fully tracked (all collaborative, and seven individual writing and reviewing processes) was performed. This analysis revealed two trends: (a) the asynchronous collaborative writing process in all groups is characterized by a constant intertwining of writing and revising (deleting, rewriting, reshuffling) activities; (b) the recursive writing style that was observed in most collaborative groups stands in stark contrast with most of the observed individual writers. The latter clearly depict a rather linear writing process, starting from the text skeleton, copying and reformulating the propositions in the given order, occasionally changing a cohesive tie or the syntax, but rarely or never changing the proposition order and/or choice.

The screenshots in the appendix illustrate two representative text geneses. Figures 1 ("Writing start in a collaborative text") and 2 ("Final edit in a collaborative text") represent
screenshots from the beginning and final phase of the writing process in one group. The different colors represent the recent changes made by different contributors working simultaneously in the document (a Google Docs built-in function). Figure 1 shows the simultaneous, but diverging, writing start of two group members. In the upper part of the screenshot, one student used the bullet list he had copied into the document in order to reformulate the facts in his own words, bit by bit deleting the bullets after having integrated the respective content in his text fragment. Meanwhile, another student started writing her fragment from scratch (in the lower part of the screenshot). Figure 2 shows the final edit carried out by one member who had been assigned by the group for this task, as recommended in the micro-script. Prior to this final edit, several revision cycles had taken place in which all group members were involved. The screenshot in Figure 3 illustrates the writing style that was identified as typical of the individual setting. The copied text skeleton serves as a model to formulate new text and is gradually deleted. The screenshot was taken in the final phase of writing the first (and final) draft, after which only minor editing activities took place.

**Quantitative and qualitative analysis of the review activities**

In order to gain insight into the collaborative review activities, the peer comments provided in the Google Docs were analyzed both quantitatively and qualitatively. The intensity of the activities varied greatly among the groups. The episodes count ranges from 3 to 34 comments. Most comments lead to a direct editing action without further discussion. Few comments were neglected. The number of comments that elicited discussions of more than two turns was also different from group to group, from 1 in the least interactive group to 15 in a group that engaged in lively interaction. The group that showed by far the highest level of review activity consisted of three highly proficient students. This group also obtained the best final score. Although there was a noticeable tendency of a higher review activity level in groups consisting of at least two highly proficient students, no significant correlation between those two variables could be detected. Table 3 shows the results of a Spearman´s Rho correlation test between final score, group proficiency level, and interaction intensity.

**TABLE 3**

Spearman´s Rho Correlations Between Final Score, Group Proficiency Level, and Interaction Intensity

<table>
<thead>
<tr>
<th></th>
<th>Group proficiency</th>
<th>Final text quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review activity</td>
<td>( r = 0.382 )</td>
<td>( r = 0.482 )</td>
</tr>
<tr>
<td></td>
<td>( p = 0.454 )</td>
<td>( p = 0.274 )</td>
</tr>
<tr>
<td>Final text quality</td>
<td>( r = 0.574 )</td>
<td>( p = 0.234 )</td>
</tr>
</tbody>
</table>

The contribution topics covered in the peer interaction were assigned to twelve categories. Following Kessler et al.’s (2012) subdivision of contributions in collaborative learning situations, nine of the twelve established topic categories refer to language-related contributions (LRCs), while the other three refer to non-language-related contributions (NLRCs): layout questions, workflow and appraisal. Table 4 shows an overview of the contribution topics in order of frequency.
There is a noticeable predominance of morphology-related discussions, covering 22.4% of the total number of contributions. Lexical choice is the only topic that was covered in all groups. Following Faigley and Witte’s (1981) taxonomy of revisions, 67% of all contributions (italicized in Table 4) can be accounted for by meaning-preserving surface changes, that is, LOCs. However, 61% of the peer feedback in the group with the best outcome involved the HOCs style and content. This group is also the only one where praise statements were included in the peer feedback.

The following excerpt from a group discussion, through Google Docs comments, illustrates a high level of reflection on language. In this discussion about a lexical choice, different arguments are presented (semantics, lexical diversification) and external authorities (two renowned bilingual dictionaries) are cited. The example also aptly illustrates the problem-solving process of the group, which is trying to establish a compromise by referring to different arguments and authorities.

Excerpt 2: Group discussion in Google Docs comments section (English translation followed by the German original, italicized)

S1: Could we find another word for this, I think Lösung would be more appropriate, although I recognize [S3]s attempt to avoid the recurrence of Lösung that was already used in the previous sentence...

S2: I used Klärung instead of Lösung in the conclusion in order to avoid recurrence.
S3: Is Klärung a proper synonym for (Auf)lösung??
S1: Indeed, I also wonder...
S1: In Duden [main unilingual German dictionary], I find Lösungsmöglichkeit, sounds somewhat abstract, but is certainly a synonym...
S2: aah, ok
S2: in Van Dale [main bilingual German-Dutch dictionary] -> een probleem tot een oplossing brengen : ein Problem einer Lösung/Klärung zuführen
S3: hmm I think that Möglichkeit is better, but if it’s 2:1 => change
S2: I also like Möglichkeit, it might be the better option.

S1: Können wir kein anderes Wort nutzen, Lösung wäre hier besser, aber ich verstehe was [S3] tun will, im vorigen Satz steht auch schon Lösung...
S2: Ich habe in meinem Schluss Klärung benutzt, anstatt Lösung, um die Wiederholung zu vermeiden
S3: Aber ist Klärung ein reiner Synonym für (Auf)lösung??
S1: ja, frage ich mich auch?
S1: In Duden finde ich "Lösungsmöglichkeit", etwas abstrakt, aber sicherlich synonym...
S2: aah, ok
S2: in 'Van Dale' -> 'een probleem tot een oplossing brengen' : 'ein Problem einer Lösung/Klärung zuführen'
S3: hmm ich meine, dass 'Möglichkeit' besser ist, aber wenn 2-1 => ändern
S2: Ich finde Möglichkeit auch gut und vielleicht besser

Attitudes Analysis

The students’ attitudes and preferences expressed in the post-hoc survey helped shed light on their perception of strengths and weaknesses of the collaborative writing and review process. Therefore, this data source will be included in the discussion of the process analysis results with a view to triangulating the qualitative observational data concerning the following aspects: (a) their experience of specific tools and aspects of the collaboration, including hindering factors, and (b) their preference toward individual or collaborative writing in this specific setting and task.

(a) 29% of the students declared their attitude toward online collaborative writing had changed for the better during the project thanks to the high usability of Google Docs. The following quote from the survey illustrates this statement.

Excerpt 3: Quote from post-hoc survey (English translation followed by the Dutch original, italicized):

Normally, I am against group work because this implies a lot of practical organization. However, this time I did not feel hindered because the group members were able to post their contributions in their own time and at their own pace.

Normaliter ben ik tegen groepswerken, omdat het veel praktische planning vraagt. Op deze manier vond ik het echter helemaal niet storend, de groepsleden kunnen op hun eigen tempo wanneer het past hun bijdrage leveren.

Collaboration in itself was perceived as fruitful, especially during the review process, as the mean rating of the following statements on a five-point Likert scale suggests (English translations from the German original): "The collaboration went well in my group" (3.27/5), "I learned from receiving peer feedback" (3.25/5), "I learned from providing peer feedback" (3.09/5), "Our feedback helped improve the final text quality" (3.71/5), and finally, "I am happy with our final text" (3.56 /5).
Nevertheless, some problems were reported. Besides free-riding, style and pace matching and the higher workload were mentioned as drawbacks of collaborative writing.

(b) When juxtaposing the two writing conditions in this specific task, students displayed two apparently diverging reactions.

Excerpt 4: Quote from post-hoc survey (English translation followed by the German original, italicized):

Peer feedback feels more personal than a sample solution.

*Man hat die Idee, Kommilitonen geben mehr persönliche Feedback als eine Musterlösung.*

On one hand, they positively pointed out the personal engagement of peer discussions in comparison to the anonymity of pre-programmed constructivist feedback, as Excerpt 4 shows. Moreover, regarding feedback handling in the individual planning condition, 48% of the students stated that they only "superficially browsed" the extra information provided in the feedback. On the other hand, 90% of the respondents stated that they preferred the individual guided planning activities to the collaborative planning in the online forum. The main problem reports related to diverse individual work pace and technical issues. Consequently, students suggested carrying out the collaborative planning face-to-face rather than online. On the whole, 70% of the respondents declared preferring the individual writing condition to the collaborative one. The following statement sums up the perceived pros and cons of both writing conditions.

Excerpt 5: Quote from post-hoc survey (English translation followed by the Dutch original, italicized):

Individual tasks are easier because you can just do it your way, even though you don’t make a lot of headway. You are limited by your own capacities and style. Group tasks are more demanding in terms of organization and are more time-consuming. Nevertheless you learn more, especially from the direct feedback you exchange with your peers. Actually, I think a combination of both is best. In the end you have to be able to do it on your own.


**DISCUSSION**

**Research question 1: Differences between the collaborative and individual syntheses with regard to text quality**

As the results of the quantitative analysis regarding text quality reveal, the individual texts differ from the collaborative ones in (a) fluency, and (b) content selection, but neither in (c) accuracy and complexity, nor in (d) cohesion and coherence. In the following, possible factors explaining these results will be discussed in the above order, also taking into consideration results of the process analysis, as and where appropriate.

**Fluency**

The significant difference in fluency is not considered important as the text production was not timed, and the number of words was limited. The greater length in the collaborative texts can be attributed to a conglomeration effect of the individual contributions.
Content Selection

Far more interesting is the finding that the collaborative texts score significantly higher on content selection than the individual texts. Groups dropped far fewer of the propositions of the source texts that were considered important for the synthesis by the teacher than individuals did. Analysis of the process data revealed that this is due to in-depth discussions during the planning phase. This means that the selected communication tool (online forum) and scaffolding method (script) for this task actually served their purpose. However, the students themselves perceived the forum as inconvenient for synchronous planning activities. Apart from problems like diverse individual work pace and technical issues, some group members disregarded the step-by-step instructions in the micro-script, which made for a rather chaotic planning phase in some groups. The danger of "over-scripting" as described by Dillenbourg (2002) might come into play here. To sum up, although collaborative online planning was perceived as inefficient by the students, it turned out to be ultimately effective.

On the contrary, the sample solutions did not cause the individual writers to change their previous own selection of propositions. This kind of canned constructivist feedback apparently did not cause the students to reflect on their own text. Perhaps they did not even examine the sample solutions in any great detail, as their responses in the post-hoc survey suggest. Why did they, then, declare to prefer individual planning? Time efficiency seems to be the key to understanding the students` attitudes. Moreover, writing habits also might play a role, as this was the first experience with online collaborative writing for most of the students.

Accuracy and Complexity

The results do not confirm claims made in previous studies that collaboration leads to higher accuracy compared with texts produced by individuals (Fernández Dobao, 2012; Storch, 2005; Wigglesworth & Storch, 2009). On a cautionary note, it should be stated that the small sample size in the present study makes it difficult to detect significant differences. Furthermore, in the above-mentioned studies, the ratios EFC/C (error-free clauses / clauses) and/or EFT/T (error-free T-units / T-unit) are used instead of, or in addition to, the E/W measure to determine the accuracy rate. However, as accuracy is not the main concern of this study, one measure was deemed sufficient to establish a trend. Another plausible explanation for this discrepant finding is the relatively high L2 proficiency among the students in the present study compared to the aforementioned studies. The proficiency level might also explain the fact that hardly any difference was found regarding balanced complexity.

Cohesion and Coherence

It is somewhat surprising 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. Apparently, most groups were able to overcome this potential problem by setting out clear agreements in the planning phase and by a thorough edit in the revision phase (both suggested in the micro-script).

Research question 2: Influence of the writing condition (individual or collaborative) on the writing process

The qualitative introspection data reveal that the writing process is highly influenced by the writing condition: while individual writers depict a rather linear writing pattern, in collaborative text production, a constant intertwining of the writing and reviewing phases was observed. This confirms the finding of previous scholarship (Fung, 2010) that collaborative writing displays features of a recursive writing process. In this way,
collaborative writing experience might help overcome the observed linear writing habit that is unfavorable in academic writing.

This recursive writing pattern fostered an intense reviewing activity with fruitful discussions on language and content-related issues in some groups. Two thirds of all peer comments during the writing and reviewing process are related to meaning-preserving surface changes, that is, LOCs. The observed predominance of morphology-related discussions can be attributed to the L2 in question, German, which is a strongly inflected language. The fact that surface changes were the major preoccupation of most groups is in line with previous research (Paulus, 1999). However, there is a tendency among the better performing groups (the groups with the highest scores on their final text) to have more contributions about style and content than the other groups. Some group discussions exhibited a keen awareness of HOCs, especially when two or more high achieving students were involved. Excerpt 2 exemplifies the high level of reflection on language in some collaborative groups, which has been described as "pooling of knowledge about language" (Storch, 2005) and is believed to be beneficial for language learning.

An interesting aside is that most groups held their discussions in the foreign language, German, although this was never explicitly suggested in the micro-script. On the one hand, this might be beneficial for their L2 communication skills training, as they employ hedges and other typical discourse markers. On the other hand, it might also prevent them from deepening the discussion due to language-related limitations.

The final survey results concerning attitudes confirm that the students were aware of the benefits of collaboration, especially when it came to reviewing their text. They were convinced that their final texts had improved thanks to the peer review activities. Although an overall superior quality of the collaborative texts in comparison with the individual texts could not be verified (see discussion of research question 1), it might still be true that some collaborative texts improved through changes in response to peer comments. In order to answer this question, a detailed analysis of the revision activities induced by peer comments would be needed, which exceeds the scope of this article.

Interestingly enough, notwithstanding the positive attitudes expressed toward the collaborative writing experience as a whole, more than two thirds of the students declared preferring the individual writing condition to the collaborative one. At first sight, this might just be a matter of habit. However, they also reported problems compromising a good collaboration, like differences in style and work pace. In one group (of seven), collaboration even seemed to have failed due to the free-riding attitude of two of the three members. Their collaboration broke down after the planning phase, and their Google Doc only contained one text skeleton. The only active group member finally decided to upload an individual text instead of a collaborative one.

CONCLUSION

The study presented in this paper was a short-term intervention under quasi-experimental conditions. It served to establish claims about possible benefits and pitfalls of collaborative L2 synthesis writing using Web 2.0 tools, compared to an individual setting. The gathered data are rich in terms of triangulation and specific in their focus on advanced L2 writing. The answers to the research questions were obtained through a mixed-method analysis.

With regard to the impact of the writing condition on the final text, two important findings should be highlighted. First, collaboratively written texts show a significant tendency toward better content selection and organization. This is caused by a higher processing depth in the synchronous collaborative planning than the individual planning guided by sample solutions. Although perceived as tedious and time-consuming by the participants, the forum discussions led to a significantly better content selection in the collaborative texts. This confirms that collaborative online planning, although perceived as inefficient by the students, turned out to be ultimately effective.

Second, no statistically significant difference between the individually and collaboratively produced texts could be detected in terms of accuracy. These results are not
in line with previous research findings on collaborative writing that attribute mainly higher accuracy to collaboratively produced texts. This might partly be attributed to the fact that the participants already had an advanced proficiency in the L2. A replication of this small-scale pilot study with more participants of a comparably high L2 command would be desirable in order to enlarge the statistical power of the tests.

With regard to the impact of the writing condition on the writing process, two main characteristics of online collaborative advanced writing have been detected. There is a tendency that online collaboration naturally stimulates recursive writing. This seems promising. Collaborative writing experience might help overcome the rather linear writing pattern observed in individual student writers which is unfavorable in academic writing. Further, instances of "pooling of knowledge about language" (Storch, 2005) that are believed to be beneficial for language learning could be observed. In addition, the pooling of cognitive resources apparent in several group discussions also proved to have a positive effect on the outcome. This is an important benefit of collaboration in a task like synthesis writing that imposes a high cognitive load on the learner. Although LOCs were indeed the major preoccupation of most groups, a keen awareness of HOCs could be observed in the discussions of the better performing groups.

Drawing upon socio-constructivist theories, the processing depth of newly acquired knowledge is known to be greater when discussion (with peers) is involved (Swain et al., 2002). Therefore, it is to be expected that the recursive writing patterns and the pooling of language knowledge present in collaborative writing have an enduring effect on individual language and/or (academic) writing skill acquisition. Future research should explore this possible effect in long-term studies in order to provide empirical evidence for the benefits of collaborative advanced writing in a foreign language, using Web 2.0 tools as a platform.

NOTES

1 CEF: Common European Framework; The CEF levels B1 and B2 correspond with the ACTFL Proficiency Guidelines levels 'Advanced low' and 'Advanced mid'.

2 This neologism translates as "enraged citizens" and refers to massive civil protests against several expensive public infrastructural projects.

REFERENCES


**APPENDIX**

Figure 1
Writing Start in a Collaborative Text (Screenshot)


Die Hauptgründe für die Proteste waren nicht nur die Milliardenhoch Baukosten, aber für das Projekt war auch die Abholzung des beliebten angrenzenden Schlossgartens mit alten Bäumen nötig. Der Bürger hatte aber kein Mitspracherecht.

- Planung des Projektes seit Jahrzehnten, endgültiger Beschluss zur Durchführung Anfang 2010
- Während des ganzen Jahres 2010 Proteste von immer mehr (zunächst hunderttausende) Bürgern gegen das Projekt
- Hauptgründe für Proteste: Baukosten in Milliardenhöhe, Abholzung des beliebten angrenzenden Schlossgartes mit alten Bäumen, kein Mitspracherecht der Bürger

- Parkschützer: Umweltverbände (u.a. BUND, Robin Wood) und Bürger, die den Schlossgarten haben wollen. Aktionen seit 2009 (u.a. Besetzung des Parks, um so das geplante Fällen der Bäume zu verhindern)
- Eskalation des Konflikts am 1.10.2010: Polizei setzt Wasserwerfer und Tränengas gegen Demonstranten ein, um Abholzung des Parks mit Gewalt durchzusetzen. Viele Verletzte
- Schlichtungsverhandlungen (mit bermittlung) zwischen Projektgegnern und Proponenten (unterschrift übereinstimmung)
- Ergebnis: Stuttgart 21 muss nachgebessert werden.


Haupttext:

Die Meinungen in Bezug auf dieses Projekt sind offensichtlich sehr verschieden und sind mit Demonstration verbunden.

Zitat: "Das war so ein gutes Gefühl der Zusammenarbeit und die Freude der Menschen, da etwas tun zu können."

Die Leute kommen zusammen aus verschiedenen Motiven. Zunächst gibt es viele Leute, die sich gegen das Projekt richten. Aber auch Leute, die sich gegen die Stadtentwicklung richten, die Art, wie die Regierung mit der Stadt und Bevölkerung umgeht. Viele protestieren eben gehen das Vorgehen der Polizei.


Stuttgart 21


ACKNOWLEDGEMENT
I wish to thank the two anonymous reviewers and the editors for their valuable comments.

AUTHOR’S BIODATA
Carola Strobl is a Research Assistant at the Faculty of Arts and Philosophy of Ghent University. She has a teaching career as a German lecturer at universities in Germany, Italy, Portugal, and Belgium, and has been involved in several research projects on technology-enhanced language learning. Currently, she is writing her dissertation entitled “Exploring the possibilities of online scaffolding for L2 academic writing in individual and collaborative settings”.

AUTHOR’S ADDRESS
carola.strobl@ugent.be