Identifying Gender Differences in CSCL Chat Conversations

Costin Chiru, Traian Rebedea, Stefan Trausan-Matu, "Politehnica" University of Bucharest, Department of Computer Science and Engineering, 313 Splaiul Independetei, Bucharest, Romania
Emails: costin.chiru@cs.pub.ro, traian.rebedea@cs.pub.ro, stefan.trausan@cs.pub.ro

Abstract: It is well known that there is a world-wide gender gap in most STEM domains. We propose a study of the participation of undergraduate students in computer science to a CSCL experiment in order to detect possible differences between female and male students. Moreover, we have tried to determine if the composition of the groups influences the value of the factors proposed for analyzing the activity of a student. The factors used for analysis are qualitative heuristics used to determine the activity of the users with regard to both involvement and content. Thus, we have been able to identify differences between the knowledge, innovation, involvement and vocabulary manifested by each gender group in several cases: chats only with males, with a majority of males or females and with an equal distribution of genders. The main conclusions of the research are that females are innovative and that equally distributed groups have higher scores than the others for several indicators.

Introduction

The gender gap in STEM (Science, Technology, Engineering, and Mathematics) studies and achievements has been widely debated during the last years and several measures for reducing this gap have been proposed both by governments and companies (Miyake et al., 2010). This is particularly true in computer science, where the number of females enlisted in undergraduate and graduate studies is far less than that of males especially in well-developed countries, but less pregnant in developing countries (Gharibyan & Gunsaulus, 2006). In order to reduce this gap, several educational strategies have been tried. For example, pair programming of women with other colleagues of the same gender has shown an increase in course completion rates (Werner, Hanks, & McDowell, 2004). However, there have been no initiatives that we know of to determine whether women involved in collaborative learning tasks enhance the performance of the groups they are part of or whether the group structure influences the performance of both women and men. These problems are tackled in the current paper as we have tried to identify the influence of the gender distribution for the individual and group outcomes in Computer Supported Collaborative Learning (CSCL) tasks involving online discussions.

Instant messenger (chat) is a main technological substrate for collaborative knowledge building in small groups (Stahl, 2006). Naturally, due to the massive extent of this type of applications, especially among young people, they have been used also for CSCL (Stahl, 2009). Probably one of the reasons for their success is their online character, implying the direct involvement of the participants. Another probable cause is their easy availability at any time and in any place. Eventually, the fact that participants do not have a face-to-face contact may induce easier communication, especially for shy or introvert persons.

However, the composition of groups and their diversity have an important impact on the collaboration accomplishment (Chidambaram & Carte, 2005). By considering various groups of different age, sex and work experience, they found out that groups that have either a perceived *or* an actual diversity had a "deeper and/or broader thinking about the problem" together with greater "idea exchange" (Chidambaram & Carte, 2005). However this was not observed for groups that had *both* an actual and perceived diversity. Other researches show that heterogeneous gender groups using computer-mediated communication exchange more messages to take a decision than homogenous groups (Savicki, Kelley, & Lingenfelter, 1996). An important aspect that goes beyond the current investigation is that groups with several leaders outperform groups with fewer leaders, but a correlation between leadership and gender has not been found (Carte, Chidambaram, & Becker, 2006).

In the case of chats and, in fact in the majority of CSCL applications, written language is the main communication media. Consequently, the success of CSCL sessions depends on a good and valuable written communication. If we take into account also the issue of group composition and diversity, we may conclude that we should investigate how different categories of participants use natural language in collaborative chats to determine possible differences between genders and caused by the gender distribution.

In this paper we focus on considering only one feature characterizing diversity in CSCL groups, the fact if the participant is a male or a female. The cognitive and affective particularities of each of the two groups of persons were studied in many researches in psychology and sociology (Eagly & Karau, 1991; Hyde, Fennema, & Lamon, 1990), but is almost absent in the case of CSCL.

Our analysis is starting from the hypothesis that males and females have different linguistic behaviors when talking in interaction and that the majority of one of the other category (or even the total absence of one) in the CSCL teams has important influences on the collaborative process and thus on the outcomes of the chat. In order to assess these outcomes, we have used a set of quantitative measures (Chiru et al., 2011) to measure knowledge, innovation, involvement and vocabulary for each participant and overall for each group.

We perform the analysis of collaboration in a corpus of CSCL chats starting from a theoretical framework that we developed in last years, based on discourse modeling in the case of online chats with multiple participants (Trausan-Matu & Rebedea, 2010). This conceptual basis was the starting point for several implemented systems (Rebedea et al., 2011), which combine Natural Language Processing (Jurafsky & Martin, 2009) with Social Network Analysis and considering also the Conversation Analysis basic ideas, such as adjacency pairs (Schegloff & Sacks, 1973). In the current case, we consider repetitions of words in subsequent utterances to be an importance factor of analysis. However, the analysis has been extended to also use lexical chains computed on the whole corpora of conversations using semantic distances computed on WordNet (Miller, 1995).

The paper continues with a description of the application developed for the comparison of chats with different gender distributions. The factors used to assess each individual and group outcomes are also presented in this section. The next section contains the main results of the comparative analysis that was performed on three levels: situation analysis for each type of group type, individual situation analysis and vocabulary analysis. The paper ends with the conclusions extracted from the previous analysis.

Designed Application

The purpose of the application developed for the research presented herein was to determine a model for the way different participants to CSCL conversations act and 'talk' in different conditions. We were interested especially in identifying if there are major differences between the participation of male and female subjects in this kind of online conversations and to decide whether the gender composition of the discussion groups may influence their outcomes.

Learning Scenario and Chat Corpus

The application was used to assess chat conversations created by senior year undergraduate students involved in a Human-Computer-Interaction (HCI) class. The students were asked to debate about different web-collaboration technologies (forums, blogs, chats, wikis, wave, etc.), highlighting the weaknesses and strengths of the existing tools and eventually devising a way to combine these tools in order to obtain an instrument that would be useful for sharing information and collaboration in a company. Each participant had to study individually the given technologies in order to identify their advantages and problems before the conversation started and afterwards they had to choose one of them and to support it in front of the other participants during the chat. The proposed scenario is similar to one of the most used macro-scripts in CSCL: the RSC (Research-Structure-Confront) script (Dillenbourg & Tchounikine, 2007). However, in our scenario the structuring stage has not been requested explicitly to the students.

The main purpose of these conversations was to ease the learning process about the considered platforms by providing each participant the possibility to critique the others' platforms and in the same time to defend its own. This way, they were able to 'see' the platforms from different perspectives and had the chance to make a comparative analysis of all the considered platforms. Finally, they had to combine the existing technologies and to develop use case scenarios so that the advantages offered by each of them to be exploited to the maximum and in the same time to eliminate the identified problems related to each individual technology.

The corpus consisted of 21 different chat conversation, each of them ranging from 158 to 579 replies, for a total of 7346 utterances.

Participant Analysis

In order to obtain the corpus of chats, we started from a group of 114 students and gave them the possibility to group in smaller teams as they wanted. The distribution considering the participants' gender was 74 males and 40 females. Five of the 74 males were involved in two different chats (were part in 2 different teams). At the same time, two females participated in two different teams. In the end, they grouped in 21 teams consisting of 4 to 11 participants, each group delivering one of the chat conversations that we analyzed (see Table 1). An important observation related to the size of the created teams is that there have been both smaller teams of 4-5 students, but also larger ones of up to 11 participants, depending on the advice the students received from their tutors regarding the advised team size: smaller teams of 4 (6 teams) or 5 (7 teams) students as compared to larger groups: 2 teams of 6, 7 or 8 participants, and 1 team of 9 or 11 participants.

The students were allowed to choose nicknames that could be used to identify them in order to be able to grade them or to have a completely anonym nicknames. Although they were assured that their (nick)names will not be made public and the usual privacy rules will be applied, some of them preferred to choose the anonymity. This is why in Table 1 there were 4 participants whose sex could not be identified based on their nicknames. These 4 cases were manually investigated by human experts and starting from their nicknames (Freaky-wiki/Wikilie, ThirdUser, Me2, BRIO), the distribution of sexes for the rest of the users, their words in the chats and their behavior, they were assumed to be males.

Table 1: Chats' length and participants' distribution

Chat No.	No. Utts.	No. Females	No. Males	No. Unknown	Total	Anonym Females	Anonym Males	Anonym + Unknown
1	203	1	3	0	4	1	1	2
2	342	3	2	0	5	3	2	5
3	289	2	3	0	5	2	3	5
4	340	1	3	1	5	1	3	5
5	311	2	2	0	4	2	2	4
6	211	1	2	2	5	1	2	5
7	397	1	2	1	4	1	0	2
8	549	1	6	0	7	0	1	1
9	347	2	7	0	9	1	1	2
10	254	3	2	0	5	1	1	2
11	158	3	3	0	6	0	1	1
12	493	0	8	0	8	0	0	0
13	541	4	4	0	8	2	0	2
14	467	2	4	0	6	1	1	2
15	203	0	5	0	5	0	0	0
16	502	7	4	0	11	2	0	2
17	463	2	5	0	7	0	0	0
18	196	1	3	0	4	1	1	2
19	579	1	3	0	4	1	3	4
20	181	3	1	0	4	3	1	4
21	320	2	3	0	5	2	3	5
All	7346	42	75	4	121	25	26	55

An interesting thing related to people's identification was that although most of the users provided enough information for establishing their gender (e.g. a first name), a lot of them chose not to provide their family name as well (this situation is defined as partially anonymised). Therefore, out of the 42 female nicknames that were identified in the chats, only 17 were providing their full names, while in the males' case 51 out of the 79 nicknames contained the participant's full name. Another interesting observation is that the most partially anonymised participants were found in small teams: 8 of the 13 teams of 4 or 5 participants consisted only on fully or partially anonymised participants, (representing 37 of the 55 anonymised participants), 4 of them had 2 partially anonymised participants and the last one had no anonymised participants. At the same time, the larger teams had at most 2 partially anonymised participants. This behavior might be explained by the need to be identified by the rest of the team: in the case of a small team, everyone knows very well the other 3 or 4 participants (so the family name is not necessary for in-group identification), while in larger teams is much more difficult to know all the members of the team well enough in order to avoid using family name. Another explanation might be given by the fact that in larger teams, there is a higher probability of having multiple persons with the same name, and therefore the family name is required for discriminate between them.

Heuristics used for participants' evaluation

In order to be able to evaluate if there are differences between the ways males and female participants act in different situations, first we had to be able to evaluate the contribution of each participant to the conversation. In order to perform this analysis, we have considered several heuristics that were previously suggested by Chiru et al. (2011):

- *Number of replies* indicating how interesting the conversation is for the considered participant;
- *Activity* of each user showing how complex one's replies are;
- *Absence* from the conversation of a participant (denoting listenership or lack of interest);
- *Persistence* of the user in the conversation, expressing the intention of a participant not to 'give the floor' to the others and computed as the number of consecutive replies issued by the participant;
- *Repetition* of other participants' concepts offering insights about how much attention a participant devoted to the content uttered by the others;
- *Usefulness* of the participant in the conversation stating how much the other participants benefited from this user's replies;
- *On topic* assessing the seriousness of the participants. It expresses how devoted the participant was to keep the conversation on the right track, considering the imposed topics for the discussion.

From the heuristics proposed by Chiru et al. (2011), we did not consider the topic rhythmicity because our purpose is somehow different from the ones of Chiru et al. (2011). We were not so interested in how often a specific topic came 'on the floor', but in building a model of how the participants were debating in different conditions. On the other hand, we investigated two different heuristic – *participant's innovation*, expressing the number of concepts introduced in the conversation by each participant and *participant's knowledge,* expressing the percent of the concepts introduced by the participant that are semantically connected with the ones imposed for debating (they are computed using a semi-automatic method of combing WordNet similarities with manual enrichment for words that have a high frequency in the analyzed corpus).

Analysis and Results

Since we were interested to see if there are differences between the ways males and females act in different situations, we considered five different scenarios: conversations with only male participants, conversations with fewer females than males, conversations with equal number of males and females, conversations where the majority of participants were females and conversations between females only. As it can be seen from Table 2, most of the chats consisted on both males and females, but with the majority of participants being males (12 chats). The lowest number of teams was observed for the conversations between participants from the same gender – only two chats between males and none involving only females.

Table 2: Distribution of chats and participants for the five classes of analyzed chats.

Distribution	Males Only	Majority Males	Equal share	Majority Females	Females Only	Total
No. Chats	2	12	3	4	0	21
Girls	0	1+2+1+1+1+1+2+2+2+1+1+2 = 17	2+3+4=9	3+3+7+3=16	0	42
Boys	8+5=13	3+3+4+4+3+6+7+4+5+3+3+3 = 48	2+3+4=9	2+2+4+1= 9	0	79

In order to analyze each class of conversations, the user has the possibility to choose one of the 5 classes or the overall statistics (see Figure 1 where the options are shown under the "Context" dropdown list).

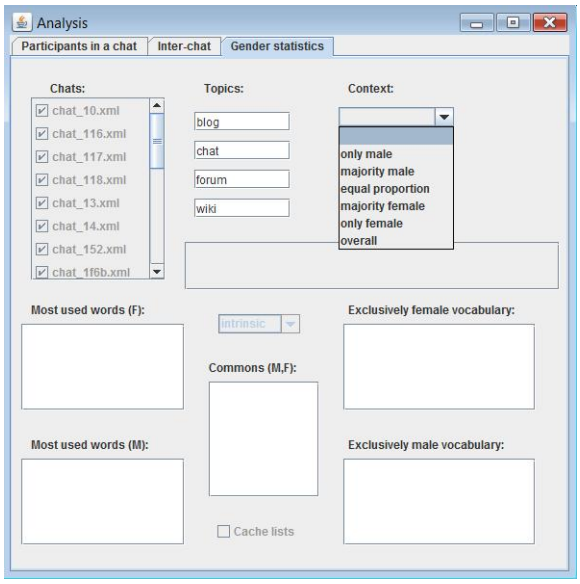

Figure 1. Front screen of the application.

Once the user has selected the desired situation, the statistics related to how male and female subjects acted in that situation are computed, and several diagrams are presented: involvement, knowledge, innovation, the percent of the vocabulary that was common and the percent of the vocabulary that is on topic. For example, see Figure 2 presenting the overall statistics. *Involvement* was computed as an average value of the scores received by the participants for the first five heuristics mentioned earlier (number of replies, activity, absence, persistence, repetition). The usefulness was used for computing how *innovative* each participant was (how many of the concepts introduced by that user were overtaken by the other participants). This heuristic, in combination with the on topic heuristic, was used to detect participants' *knowledge*. Finally the on topic heuristic used alone showed us the participants' *seriousness*. The last diagram - percentage of the vocabulary that was common – was used in order to see what share of females' vocabulary is also used by males and the same thing in the case

of males' vocabulary. In the example from Figure 2, 62.21% of females' vocabulary was also used by males, while only 41.74% of the males' vocabulary was adopted by the females, the rest being used exclusively by males.

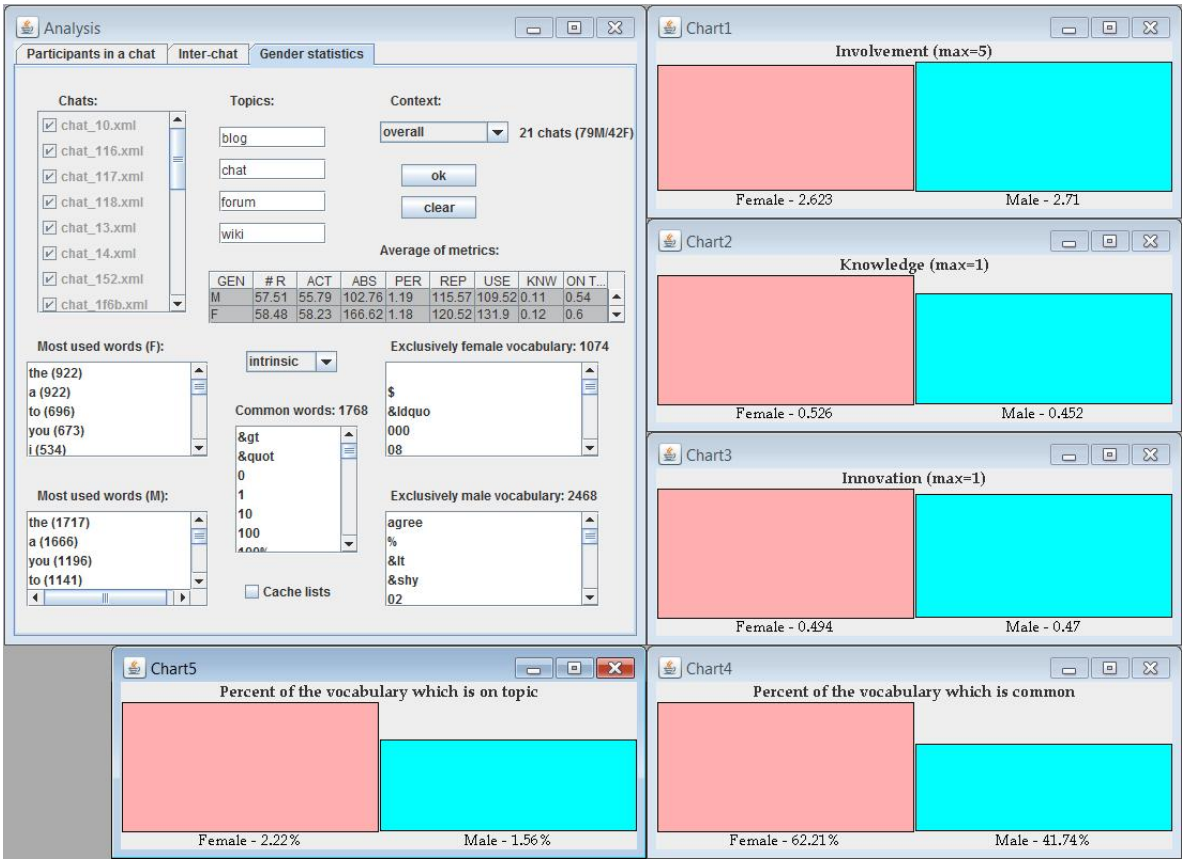

Figure 2. Application's output for the whole corpus of chats.

Since we were interested in comparing the participants acting in different situations, we have evaluated the conversations from each of the 4 different classes that we encountered (there was no chat consisting only of females) and presented the results in Table 3. Starting from these results, we can evaluate the activity of males and females in the same situations comparing to one another, or we can consider the activity of only males or females in different scenarios.

Table 3: Average scores received by the chats from the four classes identified and the overall statistics.

No chats	Situations	gen	#R	act	abs	per	rep	use	knw	on topic
21	Overall	M	57.51	55.79	102.76	1.19	115.57	109.52	0.11	0.54
		F	58.48	58.23	166.62	1.18	120.52	131.9	0.12	0.6
4	Majority Females	M	42	71.98	188.49	1.11	113.89	105.56	0.12	0.64
		F	51.69	67.27	258.13	1.13	134.94	139.62	0.14	0.65
3	Equal Share	M	53.11	50.83	108.94	1.12	98.89	84.89	0.09	0.5
		F	54.67	49.84	140.39	1.12	100.44	114.44	0.09	0.55
12	Majority Males	M	63.06	51.78	86.81	1.25	117.56	111.79	0.11	0.51
		F	66.88	54.16	94.37	1.26	117.59	133.88	0.12	0.57
2	Only Males	M	50.77	62.8	97.98	1.1	120.92	120.92	0.14	0.62

Situation Analysis

The first analysis that we did was to compare the activity of the same gender participants in different scenarios. Therefore, the males proved the least involvement in the teams formed exclusively by males (2.611) and were the most involved in the equal share teams (2.82). The males' average involvement was 2.71. Regarding the knowledge heuristic, the average was 0.452 and the most knowledgeable males were identified in the equal share teams (0.562) while the opposite was found in the case of teams formed by a female' majority (0.209)

where they probably got distracted. The innovation was also influenced by the females' participation: when they were on they own, the males were the least innovative (0.43), while the most innovation from their side was seen in the case of the female majority teams (0.492), average being 0.47. Considering the seriousness, the males' words were most of the time on topic in the female majority teams (3.95%), and least of the time in chats where there were also females, but males represented the majority (1.87%).

In the case of females, they were involved more in the cases where they did not represent the participants' majority (2.77 – equal share, 2.76 – majority males, 2.38 – majority females, 2.62 – average). They showed the highest levels of knowledge when the genders distribution was not equal (0.554 – majority males, 0.512 – majority females, 0.499 – equal share and 0.526 – average). Considering their innovation, the more females were in the team, the more innovative they were: 0.545 – majority females, 0.465 – equal share, 0.461 – majority males and 0.494 – average. Finally, the females where the least serious when they represented the majority in the chat (2.93), being most serious in the case where gender distribution was equal (3.81).

Individual Situation Analysis

The second analysis regarded the comparative behavior of males and females in each of the situations. Since in the first case there are only males, we considered only the remaining 3 cases (plus the overall statistics): majority males, equal share and majority females.

The case with majority males was characterized by very small differences in the way males and females acted considering the first five heuristics representing the participants' involvement. The females uttered a little more content (both in terms of number of utterances and of number of characters/reply). On the other hand, they were missing longer periods of time from the conversation. In the same time, they were more persistent and used slightly more repetitions. Therefore, there is almost a balance between the males' and females' involvement, with a small advantage for the females. But once we get to the qualitative evaluation, we see that there is a big difference between males and females, the latter being more knowledgeable and more serious, trying much more to keep the conversation on the right track. The males stand out only for innovation.

The second situation considered was the one when the number of males and females was equal. In this case, the females uttered the most utterances, but the males' ones were more elaborate. Again, the females were absent more time than the males from the conversations and when they 'spoke', they used more repetitions than males. In this situation, the males and females seemed to have similar knowledge, but the latter proved to be more innovative. They were also more serious than the males.

The last situation analyzed was the case when the females represented the majority in the chat. In this case it can be observed the largest gap between the average number of utterances introduced by females (51.7) and males (42). On the other hand, the males introduced much more content, using utterances much more elaborated. As in the other cases, the females were more absent than the males and used more repetitions. In this case the females proved to be more knowledgeable and innovative than the males (this case having the largest gap in both statistics between the two genders), but this time the male and female participants had almost the same score in the case of seriousness.

Finally, the overall situation presents a general statistics of all the chats that have been evaluated. This statistic is somehow biased to favor males, since there was no chat having only females as participants, while the opposite situation was found in 2 chats. Still, the results show that most of the time females seemed to be more communicative than the males (both in terms of average number of utterances and number of characters), but they were also absent for larger periods of time from the chat. They used more repetitions and proved to be more useful in the conversation. Their uttered content showed they were more knowledgeable than males, more innovative and they also proved to be more serious, trying to use words that were on topic more often than the males. Still, the males seemed to be more involved, the only heuristic where they outperformed the females.

Vocabulary Analysis

Considering the vocabularies of the 4 individual classes, along with the overall statistics, we can extract several data: the vocabulary size of the whole corpus was 5310, 2468 of the words being used exclusively by males and 1074 words being exclusively used by females (see Table 4). Most of the vocabulary was generated in the chats with majority male participants, which is normal, since this category had the most chats. If we consider the normalized vocabulary with respect to the number of chats, we can observe a mean of 252.85 words, with the highest value obtained in the case of chats between males only – 619.5, the next value being obtained in the case of equal share of males and females – 524.66. If we consider the normalized vocabulary with respect to the number of participants, the average was 25.57 words exclusively used by females and 31.24 words exclusively used by males. The average number of words used by both males and females was 14.61, but it is influenced by the fact that in the case of only males chats, the size of common used words was 0 (since there were no female participants). Most common used words are observed in the case of equal share of males and females.

Table 4: Vocabulary sizes on each of the four classes identified and the overall statistics.

Situations	Females Voc.	Males Voc.	Common Vocabulary	Vocabulary size	Norm. Females	Norm. Males	Norm. Common	Norm. Voc.
Overall	1074	2468	1768	5310	25.57	31.24	14.61	252.85
Majority females	876	440	624	1940	54.75	48.89	24.96	485
Equal share	512	550	512	1574	56.88	61.11	28.44	524.67
Majority males	604	2111	1200	3915	35.52	43.98	18.46	326.25
Only males	0	1239	0	1239	-	95.3	-	619.5

Considering each of the possible situations, we can observe that in the case of conversations with male majority, 66.52% of the female vocabulary was also used by males, while only 36.24% of the male vocabulary was used by females. These values are very close to 50% in the case of chats with equal share of males and females (50% in the case of females and 48.21% in the males' case). Finally, in the last situation (chats with female majority), the percentage of the female vocabulary that is common is 41.6%, while in the case of males is 58.65%. If we are interested in the quantity of the males' and females' vocabulary that is on topic, these values are 3.15% in the case of males-only conversations, 2.94% for females and 1.87% for males in the case of males dominated chats, 3.81% for females and 3.48% for males in the case of equal share of males and females and 2.93% for females and 3.95% for males in the case of females dominated chats. The average values for the share of on topic words that was seen in males' and females' words were 1.56% for the males and 2.22% for the females. Next, we present the top 30 words used by the males and females in each of the situations in Table 5.

Table 5: Comparison of the most frequent words appearing in the four studied cases and overall.

Only males	F	-
	M	you (195), i (132), can (130), that (95), wiki (89), chat (87), be (75), but (68), if (64), this (61), think (50), information (50), forum (49), have (48), blog (46), we (44), blogs (43)
Mainly males	F	you (304), i (216), that (180), are (126), can (121), but (110), chat (104), have (99), this (86), wiki (67), be (66), all (61), one (60), forums (59)
	M	you (740), that (569), i (556), for (432), can (325), are (296), be (283), we (262), chat (253), have (243), wiki (195), all (171), think (170), forum (165)
Equal share	F	you (93), i (74), for (61), can (57), that (52), be (47), have (40), chat (38), are (37), we (35), blogs (34), all (30), people (30), they (29), forums (28)
	M	you (108), i (83), that (68), for (60), are (57), can (53), chat (44), be (42), all (39), have (39), this (34), we (27), forums (27), good (26), one (24)
Mainly females	F	you (250), i (230), that (160), can (147), for (118), be (117), are (112), blog (109), wiki (98), chat (91), forum (80), have (76), we (76), blogs (70), yes (68), information (68)
	M	you (107), i (87), that (75), can (61), are (54), be (48), chat (43), blogs (38), forums (34), blog (32), one (32), think (31), have (28), wiki (28), your (28)
Overall	F	you (673), i (534), that (406), can (356), for (328), are (287), be (245), chat (238), have (225), blog (209), wiki (183), we (179), blogs (170), all (162), forum (160), this (155)
	M	you (1196), i (881), that (835), for (665), can (599), are (508), be (461), chat (434), we (364), have (363), wiki (337), this (282), think (275), blogs (274), all (271)

As it was expected, most of the top frequency words are determiners, prepositions, pronouns, negations and auxiliary verbs (some of them have been excluded from Table 5). Still, the main topics that were imposed for debating can be found in these top frequency words. For example in the male-only chats, we find between these words concepts such wiki, chat, information, forum, blog, blogs. In the chats that were dominated by males we see the words chat, wiki, forums being amongst the 30 most frequent words by females and chat, wiki, forum in the top 30 most frequent words by males. The equal share chats proposes words such as chat, blogs, people, forums for females most frequent words, while for males we can see the words: chat, forum, good. The last category seems the one that was the best from the content point of view, with words such as blog, wiki, chat, forum, blogs, information for females and chat, wiki, blogs for males. A similar situation can be observed in the overall statistics where chat, blog, wiki, blogs, forum where amongst the most frequent words used by females, while chat, wiki, blogs where amongst the top frequent words used by males.

Conclusions

In conclusion, gender-unbalanced chat teams decrease participants' involvement – the more balanced the teams are, the more involved are the participants. In the case of knowledge showed by the participants, they seemed to be opposite sides: the males proved to be more knowledgeable in gender-balanced teams and in teams composed of only males. At the same time, the females acted worst in the case of equal shares of males and females and acted much better in the other two remaining cases (where males seemed to be the least knowledgeable).

Regarding innovation, it seems that the best situation is to have teams where females represent the majority, since in this situation both males and females are the most innovative. Considering seriousness, the males and females acted different again, males being the most serious when the females where the least (conversations with majority females). Still, a good tradeoff seems to be given by the situation when we have equal shares of males and females, since in this case the women seemed to be the most serious while for the men this was the second best case.

A definite overall solution to the best way for gender distribution in CSCL chat tasks for small groups is difficult to reach. However, the presented results suggest that women tend to be more innovative, while men appear to discuss more on topic when they are in heterogeneous groups (only males, majority females). For each individual CSCL task, teams should be composed taking into account these factors as, in most cases, gender distribution influences the overall performance of the participants to the task.

Acknowledgments

This research presented in this paper was supported by project No.264207, ERRIC-Empowering Romanian Research on Intelligent Information Technologies (FP7-REGPOT-2010-1).

References

Carte, T., Chidambaram, L., & Becker, A. (2006). Emergent Leadership in Self-Managed Virtual Teams. *Group Decision and Negotiation, 15*(4), 323-343. doi: 10.1007/s10726-006-9045-7

Chidambaram, L., & Carte, T. (2005). *Diversity: Is There More Than Meets The Eye? A Longitudinal Study of the Impact of Technology Support on Teams with Differing Diversity.* Paper presented at the Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS '05.

Chiru, C., Cojocaru, V., Trausan-Matu, S., Rebedea, T. & Mihaila, D. (2011). Repetition and Rhythmicity Based Assessment for Chat Conversation. *ISMIS 2011, LNCS 6804*, pp 513-523.

Dillenbourg, P., & Tchounikine, P. (2007). Flexibility in macro-scripts for CSCL. *Journal of Computer Assisted Learning, 23*(1), 1-13.

Eagly, A. H., & Karau, S. J. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology, 60*(5), 685-710.

Gharibyan, H., & Gunsaulus, S. (2006). *Gender gap in computer science does not exist in one former soviet republic: results of a study*. Paper presented at the Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education, Bologna, Italy.

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155.

Jurafsky, D., & Martin, J. H. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*: Pearson Prentice Hall.

Miller, G. A. (1995). WordNet: a lexical database for English. *Commun. ACM, 38*(11), 39-41. doi: 10.1145/219717.219748

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the Gender Achievement Gap in College Science: A Classroom Study of Values Affirmation. *Science, 330*(6008), 1234-1237.

Rebedea, T., Dascalu, M., Trausan-Matu, S., Armitt, G. & Chiru, C. (2011). Automatic Assessment of Collaborative Chat Conversations with PolyCAFe. *EC-TEL 2011, LNCS 6964*, pp. 299-312.

Savicki, V., Kelley, M., & Lingenfelter, D. (1996). Gender and Group Composition in Small Task Groups using Computer-Mediated Communication. *Computers in Human Behavior, 12*(2), 209-224.

Schegloff, E. A., & Sacks, H. (1973). Opening Up Closings. *Semiotica, VIII*(4), 289-327.

Stahl, G. (2006). *Group cognition. Computer support for building collaborative knowledge*. Cambridge: MIT Press.

Stahl, G. (2009). *Studying Virtual Math Teams*. New York: Springer.

Trausan-Matu, S. & Rebedea, T. (2010). A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants. *CICLing 2010, LNCS 6008*, Springer, 2010, pp. 354-363

Werner, L. L., Hanks, B., & McDowell, C. (2004). Pair-programming helps female computer science students. *J. Educ. Resour. Comput., 4*(1), 4. doi: 10.1145/1060071.1060075