Emo Dialogue: Differences in male and female ways of communicating with affective autonomous conversational systems

Brigitte Krenn, Stephanie Schreitter

Background

Within the project CyberEmotions, a series of experiments was carried out in an online web chat environment to collect data on verbal behavior when exposed to affective utterances. In the context of CyberEmotions, the focus of research is how affect spread in virtual communities such as chats, discussion forums etc. The project, however, does not touch gender-related research questions.

The raw data collected in CyberEmotions were made available to C4U, where they were assessed with respect to differences between female and male dialogue partners when communicating with the artificial systems. The investigations we carried out are parallel to those presented in deliverable E3_1 and E3_2.1. Thus we can draw further evidence for differences or similarities in female and male dialogues with artificial agents.

Experimental setting

An autonomous dialog system was equipped with three affective profiles: a positive, a negative, and a neutral one. Users participating in this experiment chatted with all three variants of the dialog system. The experiment was conducted in cooperation with a Polish internet portal. The participants were recruited from their user pool.

91 Polish participants (33 female, 58 male), aged between 18 and 52, took part in the experiment. Each participant completed interactions in all three experimental settings, i.e. talking to an artificial communication partner which is in a neutral, positive or negative "mood".

The dialogues with the autonomous system were conducted in English. Even though the mother tongue of the participants was Polish, all participants had at least average communication skills in English. For more details on the data collection experiment see (Skowron2011).

Data analysis

In the analysis presented in this paper, we investigate differences in male and female ways of communicating with the three different variants of the artificial dialog systems along the following lines:

1. human assessment of the different versions of the conversational system (questionnaire)
2. dialogue analysis (number of utterances, words, characters, etc.)
3. word-level investigation using the Linguistic Inquiry and Word Count Approach (LIWC)

Questionnaire

Following each interaction, the participants answered 7 questions (on a Likert-scale ranging from 1 -- not at all to 5 --very much) whether the dialogues seemed real, coherent, positive or negative to the user and if he or she enjoyed the conversation, felt an emotional connection and would like to chat again with that system.

The questions are as follows:

1. *I enjoyed chatting with the conversational partner during the just completed interaction.*
2. *I found a kind of "emotional connection" between myself and the conversational partner.*
3. *I found the dialog with the conversational partner to be realistic.*
4. *I found the dialog to be coherent. In other words, the sequence of responses of the conversational partner made sense.*
5. *I noticed a positive emotional change in myself during the interaction.*
6. *I noticed a negative emotional change in myself during the interaction.*
7. *I would like to chat again with this particular conversational partner in the future.*

In Figure 1, the mean scores of male and female users are plotted. A Mann Whitney U test was conducted to test for significant differences between male and female evaluations of the conversations. Table 1 shows that male users felt a significantly higher *emotional connection* with the neutral (1) and the negative (3) systems. Moreover, the male participants rated the negative system as significantly more *coherent* and *realistic* than the female participants did. In all other aspects, the evaluations of female and male users did not significantly differ.

The plots in Figure 1 also show that both male and female users enjoyed communicating with the positive system most, found it least negative, and wanted to communicate again with the system, whereas the negative system was consistently evaluated as negative.

Figure 1:Line diagram of mean scores for each affective experimental setting (system 1 [neutral], system 2 [positive], system 3 [negative]) for male and female users. The y-axis indicates the participant's agreement on a Likert-scale from 1 *not at all* to 5 *very much.*

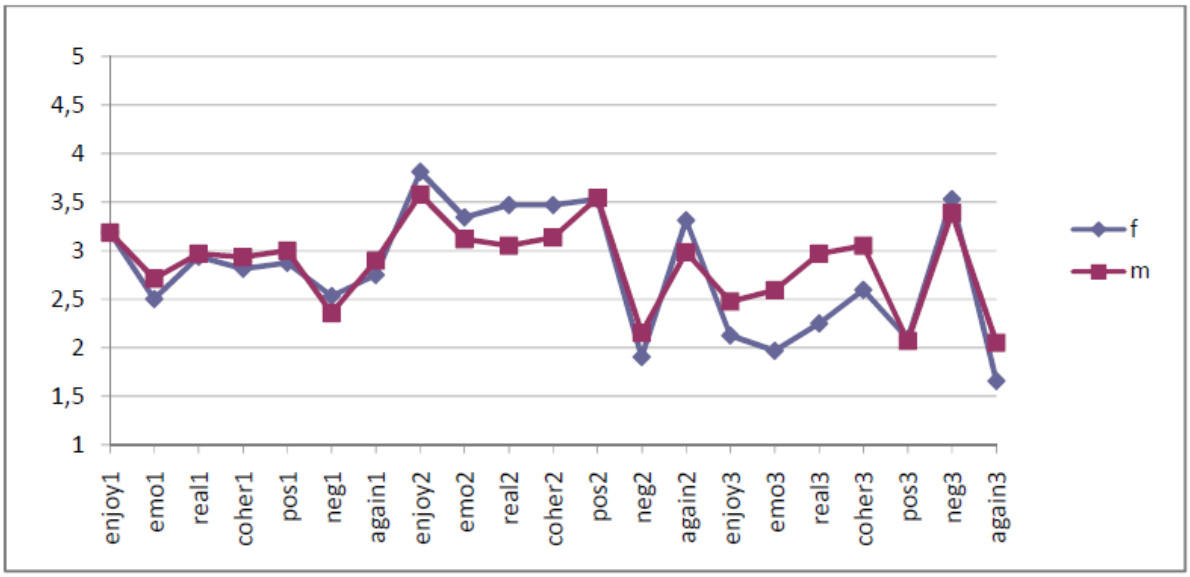

Table 1: The mean scores of the answers of male and female user and the p-values (grey indicates significance)

	Mean values		P-values
	f	m	
enjoy1	3.188	3.186	0.202
emo1	2.500	2.712	0.045
real1	2.938	2.966	0.194
coher1	2.813	2.932	0.536
pos1	2.875	3.000	0.082
neg1	2.531	2.356	0.417
again1	2.750	2.898	0.098
enjoy2	3.813	3.576	0.672
emo2	3.344	3.119	0.937
real2	3.469	3.051	0.400
coher2	3.469	3.136	0.242
pos2	3.531	3.542	0.570
neg2	1.906	2.153	0.243
again2	3.313	2.983	0.830
enjoy3	2.125	2.475	0.059
emo3	1.969	2.593	0.015
real3	2.250	2.966	0.002
coher3	2.594	3.051	0.024
pos3	2.094	2.068	0.764
neg3	3.531	3.390	0.542
again3	1.656	2.051	0.054

Dialog analysis

An analysis of the contributions to the dialogs (Table 2) in terms of number of utterances in a dialogue, number of words and characters, and dialogue act classes showed significant differences in female and male reactions to the positive and the negative systems, namely

- male dialogue partners
  - used significantly more *words* during the conversation with the positive system,
  - asked the positive system significantly *more yes-no-questions*,
  - uttered significantly more *statements* to the negative system;
- female dialogue partners

- directed significantly more *rejections* of the statements made by the negative system,
- gave significantly more *no-answers* to the negative system.

Table 2: The mean scores of the dialog act analysis of male and female users and the p-values (grey indicates significance).

Dialog Analysis

	System 1 (Neutr)			System 2 (Pos)			System 3 (Neg)		
	mean values		p values	mean values		p values	mean values		p values
	female	male		female	male		female	male	
Utterances	16.938	15.797	0.901	15.750	16.424	0.275	14.813	15.864	0.585
Words	77.938	86.288	0.173	70.781	86.847	0.025	70.813	84.780	0.200
Characters	377.344	425.220	0.122	333.250	426.475	0.011	336.438	413.288	0.141
posEmoticon	0.844	0.780	0.138	1.688	1.475	0.069	0.563	0.898	0.822
negEmoticon	0.000	0.000		0.094	0.034	0.234	0.031	0.034	0.946
Statement	0.542	0.560	0.907	0.612	0.613	0.865	0.436	0.555	0.001
whQuestion	0.000	0.000		0.000	0.000		0.000	0.000	
Reject	0.022	0.023	0.572	0.027	0.021	0.342	0.045	0.022	0.031
Emotion	0.008	0.002	0.192	0.017	0.015	0.650	0.009	0.012	0.386
Accept	0.031	0.046	0.461	0.049	0.033	0.066	0.036	0.031	0.733
Order	0.033	0.05	0.345	0.034	0.048	0.296	0.026	0.032	0.334
ynQuestion	0.062	0.058	0.529	0.029	0.058	0.034	0.087	0.071	0.306
Continuer	0.062	0.058	0.505	0.036	0.033	0.747	0.075	0.068	0.582
Bye	0.014	0.004	0.817	0.003	0.003	0.691	0.016	0.012	0.764
Emphasis	0.007	0.017	0.460	0.015	0.018	0.772	0.024	0.020	0.707
Greet	0.014	0.005	0.500	0.011	0.009	0.977	0.043	0.006	0.214
yAnswer	0.079	0.056	0.152	0.057	0.041	0.173	0.047	0.034	0.370
nAnswer	0.036	0.046	0.898	0.046	0.032	0.303	0.056	0.031	0.038

Analysis based on LIWC categories

The Linguistic Inquery and Word Count categories were developed by (Pennebaker2001).

An analysis of the LIWC categories (see Table 5) shows that during chatting with the **neutral system** female participants used significantly more *verbs*, and *present* tense as well as *health* categories than male participants, who on the other hand used significantly more *prepositions* and *hearing* categories.

Communicating with the **positive system**, women used *I* significantly more often than male did. They also agreed (*assent*) significantly more often with the system and mentioned *sexual* aspects more often, while men talked significantly more about *leisure* and used more words from the category *discrepancy.*

The dialogue contributions of male and female users differed most when interacting with the **negative system**. Female participants used significantly more *function words*, *I*, *auxiliary verbs*, *adverbs* and *negations*. Male participants used significantly more *prepositions*, and words belonging to the LIWC categories *space* and *achievement*.

For the full list of LIWC categories see http://www.liwc.net/descriptiontable1.php. The categories for which we found significant differences in female and male dialogue contributions are listed in Table 3 below. The lexical examples presented in Table 4 stem from the LIWC category table available under the URL above.

Table 3: LIWC categories for which significant differences were found.

Neutr. System		**Pos. System**		**Neg. System**	
f	**m**	**f**	**m**	**f**	**m**
+ Verbs	+ Hear	+ I	+ Discrep	+ I	+ Space
+ Present	+ Prep	+ Sexual	+ Leisure	+ Funct	+ Prep
+ Health		+ Assent		+ Negate	+ Achiev
				+ AuxVB	
				+ Adverb	

Table 4: LIWC categories with significantly different occurrences in dialogue contributions by females and males.

Linguistic processes: Adverb: adverbs, *very, really, quickly* AuxVB: auxiliary verbs, *am, will, have* Funct: function words I: 1st person singular, *I, me, mine* Negate: negations, *no, not, never* Prep: prepositions, t*o, with, above* Present: present tense, *is, does, hear* Verbs: main verb, *walk, went, see*	**Psychological processes:** Discrep: discrepancy, *should, would, could* Hear: *listen, hearing* Health: *clinic, flu, pill* Sexual*: horny, love, incest* Achiev: achievement, *earn, hero, win* Space: special words (psychological processes), *down, in, thin*
Personal concerns: Leisure: *cook, chat, movie*	**Spoken categories:** Assent: *agree, OK, yes*

Table 5: The mean scores of the LIWC categories of male and female users and the p-values (grey indicates significance).

	LIWC										
	Neutr System				Pos System				Neg System		
	mean		p-values		mean		p-values		mean		p-values
	f	m			f	m			f	m	
Funct1	0,485	0,481	0,904	Funct2	0,511	0,500	0,699	Funct3	0,535	0,512	0,040
Pronoun1	0,174	0,171	0,739	Pronoun2	0,187	0,183	0,888	Pronoun3	0,203	0,197	0,392
Ppron1	0,123	0,120	0,878	Ppron2	0,140	0,124	0,057	Ppron3	0,154	0,138	0,054
I1	0,094	0,089	0,522	I2	0,101	0,083	0,022	I3	0,086	0,075	0,029
We1	0,002	0,003	0,738	We2	0,008	0,004	0,404	We3	0,015	0,013	0,710
You1	0,024	0,023	0,732	You2	0,031	0,033	0,414	You3	0,051	0,047	0,485
SheHe1	0,000	0,001	0,494	SheHe2	0,004	0,001	0,219	SheHe3	0,002	0,002	0,621
They1	0,004	0,006	0,189	They2	0,005	0,006	0,542	They3	0,006	0,007	0,355
Ipron1	0,051	0,051	0,848	Ipron2	0,050	0,059	0,147	Ipron3	0,050	0,061	0,130
Article1	0,020	0,022	0,940	Article2	0,027	0,027	0,851	Article3	0,018	0,029	0,284
Verbs1	0,196	0,173	0,005	Verbs2	0,189	0,179	0,575	Verbs3	0,193	0,188	0,475
AuxVb1	0,112	0,098	0,098	AuxVb2	0,118	0,105	0,141	AuxVb3	0,129	0,113	0,045
Past1	0,013	0,015	0,633	Past2	0,028	0,028	0,768	Past3	0,016	0,018	0,869
Present1	0,164	0,140	0,018	Present2	0,144	0,132	0,188	Present3	0,164	0,155	0,212
Future1	0,011	0,013	0,707	Future2	0,016	0,013	0,807	Future3	0,011	0,012	0,309
Adverbs1	0,051	0,051	0,937	Adverbs2	0,060	0,054	0,439	Adverbs3	0,075	0,053	0,049
Prep1	0,059	0,074	0,016	Prep2	0,063	0,071	0,251	Prep3	0,059	0,073	0,030
Conj1	0,047	0,043	0,861	Conj2	0,040	0,040	0,752	Conj3	0,041	0,043	0,970
Negate1	0,059	0,055	0,255	Negate2	0,063	0,049	0,110	Negate3	0,074	0,054	0,002
Quant1	0,023	0,022	0,688	Quant2	0,025	0,023	0,809	Quant3	0,017	0,020	0,343
Numbers1	0,003	0,003	0,726	Numbers2	0,004	0,004	0,740	Numbers3	0,002	0,005	0,207
Swear1	0,001	0,002	0,305	Swear2	0,005	0,002	0,886	Swear3	0,004	0,008	0,508
Social1	0,066	0,063	0,777	Social2	0,077	0,074	0,917	Social3	0,131	0,109	0,201
Family1	0,001	0,002	0,533	Family2	0,003	0,001	0,097	Family3	0,002	0,001	0,196
Friends1	0,006	0,004	0,146	Friends2	0,005	0,003	0,861	Friends3	0,002	0,002	0,707
Humans1	0,005	0,003	0,432	Humans2	0,007	0,003	0,194	Humans3	0,005	0,006	0,441
Affect1	0,110	0,098	0,086	Affect2	0,109	0,095	0,113	Affect3	0,073	0,074	0,705
Posemo1	0,092	0,084	0,128	Posemo2	0,091	0,078	0,073	Posemo3	0,054	0,053	0,793
Negemo1	0,018	0,014	0,206	Negemo2	0,020	0,016	0,264	Negemo3	0,020	0,023	0,917
Anx1	0,003	0,002	0,693	Anx2	0,004	0,002	0,712	Anx3	0,003	0,003	0,944
Anger1	0,008	0,005	0,459	Anger2	0,011	0,007	0,146	Anger3	0,008	0,014	0,155
Sad1	0,002	0,002	0,416	Sad2	0,004	0,002	0,652	Sad3	0,002	0,003	0,875
CogMech1	0,154	0,144	0,383	CogMech2	0,138	0,147	0,262	CogMech3	0,159	0,155	0,530
Insight1	0,026	0,022	0,883	Insight2	0,030	0,028	0,644	Insight3	0,026	0,024	0,845
Cause1	0,010	0,016	0,060	Cause2	0,015	0,016	0,479	Cause3	0,012	0,016	0,443
Discrep1	0,016	0,016	0,618	Discrep2	0,011	0,019	0,007	Discrep3	0,023	0,027	0,222
Tentat1	0,030	0,025	0,635	Tentat2	0,024	0,027	0,626	Tentat3	0,029	0,030	0,466
Certain1	0,011	0,012	0,657	Certain2	0,013	0,014	0,816	Certain3	0,014	0,011	0,535
Inhib1	0,008	0,005	0,137	Inhib2	0,004	0,002	0,754	Inhib3	0,004	0,003	0,875
Incl1	0,026	0,024	0,523	Incl2	0,026	0,017	0,119	Incl3	0,032	0,027	0,427
Excl1	0,037	0,034	0,777	Excl2	0,038	0,039	0,674	Excl3	0,040	0,036	0,600
Percept1	0,012	0,016	0,193	Percept2	0,018	0,020	0,960	Percept3	0,012	0,016	0,319
See1	0,005	0,006	0,354	See2	0,006	0,006	0,636	See3	0,005	0,006	0,615
Hear1	0,002	0,006	0,038	Hear2	0,010	0,009	0,873	Hear3	0,006	0,008	0,394
Feel1	0,004	0,003	0,711	Feel2	0,006	0,004	0,585	Feel3	0,003	0,003	0,778
Bio1	0,047	0,052	0,311	Bio2	0,027	0,029	0,690	Bio3	0,022	0,025	0,463
Body1	0,003	0,004	0,105	Body2	0,005	0,003	0,960	Body3	0,004	0,005	0,130
Health1	0,007	0,004	0,033	Health2	0,005	0,003	0,449	Health3	0,002	0,002	0,876
Sexual1	0,001	0,003	0,133	Sexual2	0,005	0,001	0,012	Sexual3	0,003	0,004	0,437
Ingest1	0,037	0,044	0,233	Ingest2	0,020	0,025	0,318	Ingest3	0,017	0,018	0,887
Relativ1	0,069	0,066	0,730	Relativ2	0,065	0,064	0,755	Relativ3	0,062	0,059	0,211
Motion1	0,006	0,006	0,768	Motion2	0,009	0,008	0,709	Motion3	0,009	0,011	0,631
Space1	0,034	0,033	0,569	Space2	0,028	0,028	0,887	Space3	0,016	0,026	0,023
Time1	0,045	0,037	0,262	Time2	0,048	0,039	0,202	Time3	0,051	0,036	0,917
Work1	0,008	0,014	0,080	Work2	0,019	0,021	0,745	Work3	0,013	0,013	0,643
Achiev1	0,007	0,007	0,940	Achiev2	0,007	0,009	0,241	Achiev3	0,008	0,013	0,005
Leisure1	0,024	0,029	0,497	Leisure2	0,020	0,030	0,036	Leisure3	0,024	0,021	0,563
Home1	0,003	0,003	0,793	Home2	0,005	0,003	0,126	Home3	0,002	0,001	0,892
Money1	0,005	0,005	0,723	Money2	0,020	0,017	0,335	Money3	0,014	0,018	0,180
Relig1	0,006	0,005	0,681	Relig2	0,008	0,005	0,255	Relig3	0,005	0,005	0,702
Death1	0,000	0,001	0,494	Death2	0,003	0,002	0,837	Death3	0,001	0,002	0,531
Assent1	0,031	0,022	0,101	Assent2	0,032	0,020	0,014	Assent3	0,019	0,012	0,191
Nonflu1	0,005	0,002	0,244	Nonflu2	0,005	0,003	0,497	Nonflu3	0,003	0,003	0,783
Filler1	0,015	0,012	0,120	Filler2	0,018	0,011	0,249	Filler3	0,010	0,011	0,456

References

Pennebaker, J. W., Francis, M. E., & Booth, R. K. (2001). Linguistic inquery and word count: LIWC 2001. Erlbaum Publishers.

Skowron, M., Priker, H., Rank, S., Paltoglou, G., Ahn, J., & Gobron, S. (2011). No peanuts! Affective cues for the virtual bartender. In Proceedings of the Florida artificial intelligence research society conference.