Emoticons Signal Expertise in Technical Web Forums

Liliana Mamani Sanchez and Carl Vogel

Computational Linguistics Group, Trinity College Dublin
{mamanisl,vogel}@tcd.ie

Abstract. Past research has demonstrated intercultural differences in emoticon use with effects of the topic of discourse (e.g. science vs. politics) interacting with the culture of online postings (e.g. UK, Italy, Sweden, Germany). The current research focuses within a discourse, and within a lingua franca for communication and attempts to assess whether emoticon use varies as a function of user-type within the online context. The online context is a web user-forum associated with a software technology company. The user categories are determined by a few orthogonal classifications: employees, novice users, and experts; recipients of kudos vs. non-recipients of kudos; etc. As part of a developing theory of presentation of "professional" selves, and perceptions thereof, we test the hypotheses that kudo recipients deploy markedly fewer negative emoticons than comparison categories and that non-employee experts use markedly more emoticons in general than other categories of forum users. Also interactivity across the different group of users and their correlation with emoticon use was explored.

1 Introduction

Community forums are being increasingly used by companies as they provide a channel of communication with consumers. In this paper we explore the forum of a major software company in which participants discuss technical aspects of products and services. Emoticons have been used in sentiment analysis systems as clues for determining sentiment scores ([6]), for training sentiment classifiers with domain independence [7] and for collecting training data for sentiment classification in micro-blogging systems [1]. Cross-cultural analyses of emoticon use are also available [5,8]. Here, we compare emoticons and smilies usage in terms of their relative frequency across three group of forum users.[1] The distinction between emoticons and smilies has emerged over time with the potential for graphical user interfaces to depict them as composed from typographical characters (e.g. ":-)") or pictorally (e.g. ). In an example of synonymy avoidance in natural language [2], the former have become known as "emoticons", and the

[1] The discussion in the forum analyzed tends to be monolingually English, although, parallel forums exist.

B. Apolloni et al. (Eds.): *Neural Nets and Surroundings*, SIST 19, pp. 415–425.
DOI: 10.1007/978-3-642-35467-0_41

latter as "smilies"; this terminology is adopted in the present work. The forum users share the same language but have different levels of expertise, at least in a nominal dimension (all in a forum may ask or respond to questions that presuppose varying levels of expertise with the subject matter at hand, but are classified in general within the system – a participant may be given the nominal label "guru" without each posting illustrating this status). We are aware that users in forums are likely conscious of how they are perceived, and desire to be deemed as professional and expert as the employees who make contributions. Thus, we anticipate a kind of convergence on linguistic and nonlinguistic features in postings of employees as the amateur contributors participate (cf [4]).

In principle, we distinguish three groups of forum contributors, since they represent distinct aspects of a consumer related forum scenario. Firstly, common users (consumers) approach the forum in search of solutions to technical issues arising from the use of software products.[2] Secondly, gurus are facilitator-users who may have been common users and were given this "badge" name because their degree of contribution to the forum in terms of quality of answers, level of engagement and level of knowledge they share with less expert users. Gurus do not receive any financial reward for their services. The reward for them has more a subjective character for which the main motive for their contribution appears to be the prestige they can obtain, partly via feedback from other users who may reward postings with "kudos" (by clicking on this nominal label in evaluating a posting). Theories from social psychology giving accounts of this kind of dynamic have been broadly described in [3]. Thirdly, employees who are current or past workers: in contrast to the others, they may receive financial reward for their contributions to the forum, i.e. as part of their position duties. Additional details about our method are described in Section 2.

We explore the usage of emoticons and smilies as signals of emotion across users categories. The set of emoticons and smilies will be further described in Section 3 along with the dataset we used. We wanted to know the extent to which emoticon and smilie usage is related to this user classification and whether this usage was influenced by the prestige users obtain and their level of interactivity in the forum. We use the notion of how many kudos (i.e. positive ratings) posts are given as a clue of user prestige. Additional clues for determining level of interactivity we use are: the ratio of the number of posters in a group with relation to the total of posts, and the depth of readership, evidenced by the depth that a post has in a thread. These metrics are detailed in Section 3.2. We preset our results in Section 4 and finally we conclude with a discussion about our main findings and future work in Section 5.

2 Method

We decided to split the common users group between ranked and unranked users (aka, "not-ranked"). Rank is assigned to common users who have started to show levels of contribution, but differently from the guru role, they can climb

[2] We ignore short and long-term lurkers.

a hierarchy of ranks according to certain quantitative thresholds. Gurus on the other hand are assigned this role in a more subjective manner, taking into account quantitative and qualitative parameters. Therefore, the first categorization based in roles includes: {guru, employee, ranked, not-ranked} users. The second categorization is based on whether a user has received kudos: {kudoer user, non-kudoer user}. The third categorization groups posts that received kudos in one category and post that did not receive kudos in another one: {kudoer post, non-kudoer post}. Surveying over the role-based categories allows us to explore emotion expression in groups where the distinction based in levels of expertise is more fine-grained as it is implemented by forum moderators. On the other hand the kudo-based categorizations allows us to work over another dimension of user and post classification, as kudos are given by any kind of user to posts they find useful or outstanding.

We decide to work with emoticons (based on ASCII characters) and smilies (based on pictures) as they both convey emotions and we evaluate our metrics individually or jointly on both types of signals of emotion. The metrics of interactivity we evaluate are related to the volume of posts each user has created. Aspects of volume of posts are the quantity and length. These metrics were inspired by previous research done over use of emoticons based on characters in newsgroups in various languages discussing politics ([5,8]).

3 Data and Processing

The forum data was obtained from an academic alliance with the R&D department of the software company whose forums we've analyzed. The provided data comprises: actual post content in XML and HTML format (subject and body) and metadata about posts and users. The metadata around posts comprises: posting date, user id, thread id, post id, last edition time, last edition author, kudos received, and views received. Metadata about user comprises: roles and date of promotion to the guru role in case such a role applies. Body of messages comprise the authors' writing, but can also include two elements that are not written by the author: quotes referring to previous posts which are embedded in the text, either in part or entirely, and an edition field.

We used a set of 98 emoticons selected from previous research ([5,8]). Additionally, we included a set of 45 smilies (e.g.  and ) that were provided by the forum management system. Note that the smilies cover the same sorts of expressions of affect as the emoticons (e.g. smiles versus frowns), but also purport to encode additional distinctions as well (such as gender:  vs. ).

3.1 Treatment of the Data

The raw forum dataset consisted of 308,274 posts covering all the messages posted between the creation of the forum and 2010-10-12T11:24:16+00:00. We excluded posts and threads where only one or two groups of users could participate, as well as posts from users belonging to the guru group before they were assigned this role. In related ongoing work, we examine the factors that lead to such individuals ultimately becoming classified as gurus. The working dataset has 208,284 posts. After isolating the text written by a single user, we use regular expressions to count the occurrences of emoticons. Matching true positives proved to be difficult since web text often contains symbols that can be confounded with emoticons resulting in false positives such as the cases shown in Fig. 1. We chose a conservative approach and only matched emoticons well delimited by spaces, tabs or newlines characters.

FAT partition of (C**:)** and the old (F**:)** is now (E**:)**.... The storage Destinatioon was a Slave Drive (F**:)** ...
... 7) Selected Recover My Computer. **8)** In the drop-list-box above the list of recovery points, I selected ...
Can add URL or use mask (e.g. with ? or ***)**

Fig. 1. Cases of false matches for the emoticons **:)**, **8)** and ***)**

3.2 Forum Data Profile

In this section we explain the metric we used to measure interactivity. Later calculated values will be used to explore a correlation between usage of signals of emotion and level of interactivity. Table 1 shows the number of posts, users and average ratio of posts per individual (APPI) across groups for the three categorisations. The APPI gives a sense of how much interactivity is present in each group: Guru users show the highest level of interactivity, ranked users show the second highest ratio of posts, employees show less interactivity than these two groups, and not-ranked users show the least traffic of posts (almost 1 per user). Kudoer users have a high ratio of posts when compared to non-kudoer users. For the third categorisation, the number of users who authored each kind of post was counted; users from kudoer posts correspond with the number of kudoer users, while the number of users for nonkudoer posts comprises nonkudoer users plus users who have at least one post without kudos. From these figures, it can be seen that 2007 kudoer users have also some nonkudoer posts and 295 users have only kudoer posts. The APPI counts for kudoer and nonkudoer posts are not significantly different, because 2007 of the 2302 kukoer users also contributed posts that received no kudos.

Although illustrative, the APPI is a general measure of interactivity since there is no way to determine whether users have read each other's posts even if the posts belong to the same thread. In some posts, authors included quotes from previous posts, but this is not always the case. However, not including quotes does not mean the poster has not read previous posts since we consider it a default principle every reply to the thread is related to the post that started the thread.[3] These caveats understood, we explore two extra measures related to volume of messages per thread: depth of a message and thread creation.

[3] Even here, it is possible for users to contribute "off-thread" posts to a thread.

Table 1. Posts per group for the three categorisations

	Role	numposts	numusers	APPI
By role	employee	25,490	400	63.725
	guru	27,489	15	1,832.60
	notranked	50,456	19,462	2.593
	ranked	104,849	2,273	46.128
User Cat.	kudoer	139,164	2,302	60.456
	nonkudoer	69,120	19,848	3.482
Post Cat.	kudoer	18,540	2,302	8.054
	nonkudoer	189,744	21,855	8.682

Table 2. Average depth of postings for the three categorisations

	Role	Mean	sd
By role	employee	12.96	31.11
	guru	9.75	18.48
	notranked	18.49	51.14
	ranked	16.46	34.84
User Cat.	kudoer	14.35	33.15
	nonkudoer	17.94	43.47
Post Cat.	kudoer	15.45	39.26
	nonkudoer	15.34	36.05

The depth of a message in a thread signals levels of interactivity if we follow the logic that a user posting the n-th message in a thread has read the previous $n - 1$ posts. We assigned the first post in a thread a 0-depth, the next 1 and so on. This intuition that a poster of a message at level n has read all the previous $n - 1$ messages may, however, be wrong, especially cases where threads have a high number of posts. Yet, the measure gives us an idea of the latency of threads in terms of interactivity as it shows a user's tendency to participate in extant conversations. Table 2 shows the average and standard deviation of posts' depths, for the three user classifications. The less experienced users: Not-ranked, non-kudoers have the highest average of post's depth, while the depth averages for kudoer posts and non-kudoer posts are not significantly different.

The next measure we explore is whether a post is a new posting (creates a new thread) or a reply (contributes to an existing thread). Posting new messages is less interactive than replying, as a reply-poster responds at least one previous post, while thread-creators at most invite reply, whether or not their new posts engage with those of separate prior threads. Table 3 shows the number of new posts in contrast to the number of replies and the ratio of new posts to replies (NPRR) across groups.[4] The numbers show that even posting a high number of replies, non-ranked, non-kudoers and non-kudoer posts' authors write more new posts than the rest of users. Furthermore gurus are less likely to start a thread, but they are highly interactive with 75 replies in average for each new post.

On the basis of these measures, there is every likelihood that the group of gurus is the most interactive one, while for the kudo-categorisations, kudoers and kudoer-post authors are the most interactive ones. Long threads point to high interactivity because they reflect group discussion rather than bipersonal conversations. However is not possible to measure the length of a thread per group in our scenario, since threads are not exclusive to a specific group.

Table 4 shows averages of word-level tokens per post, that could help to find out levels of interactivity if we hypothesise that the most interactive users tend to write more content that could include a high use of emoticons, however this

[4] In this case, we use a version of the dataset that includes posts made by gurus before their promotion, for sake of integrity in the calculation of NPRR.

Table 3. Ratio of new posts to replies per role

	Role	NewPosts	Replies	NPRR
By role	employee	1,954	23,563	0.08
	guru	637	48,109	0.01
	notranked	16,729	33,730	0.50
	ranked	11,868	93,186	0.13
User Cat.	kudoer	11,990	150,477	0.08
	nonkudoer	19,198	48,111	0.40
Post Cat.	kudoer	1,599	20,254	0.08
	nonkudoer	29,589	178,334	0.17

Table 4. Average of tokens per post

	Roles	Tokens
By role	employee	58.63
	guru	70.70
	notranked	89.23
	ranked	74.03
User Cat.	kudoer	70.24
	nonkudoer	85.75
Post Cat.	kudoer	99.38
	nonkudoer	73.04

intuition is not correct in all the cases since sometimes users write their posts copying/pasting logs from software tools. As we pointed in Sect. 3.1 we tried to identify these elements to reduce them to one token, but it was not possible in every case. Still from these tables we can deduce that not-ranked, non-kudoer users and kudoer-post authors are the most prolific regarding post size.

That is, in this section we have profiled user categories according to proxy measures of interactivity and forum-related expertise. We conjecture that use of signals of affect, whether with emoticons or smilies, is a function of user-interactivity and expertise. We expect the more expert users to lean towards signals of positive affect, and greater levels of affect signalling with greater levels of interactivity. In next section we relate the general findings in this section to frequencies of signals of emotion usage.

4 Results

A table of the 10 most frequently used emoticons and smilies in each of the three categories of affect considered is provided in Table 5.

4.1 Usage of Signals of Emotion Across Groups

We explored how frequently signals of emotion are used in relation to the amount of messages posted by each group. The number of posts with and without signals of emotion and the proportion of use of signals are shown in Table 6. In general the use of either emoticon or smilies is low in our dataset ($\sim 6\%$) across the three categorizations, while the combined counts (emoticons plus smilies) show that at most 10% of the posts use at least one signal of emotion when the authors are ranked users. Combined proportions from kudoers and non-kudoers are not very different, altough kudoers use signals of emotion more frequently.

Posts from gurus, not-ranked, non-kudoer users and non-kudoer-posts authors include more emoticons than smilies when compared to the ones authored by employees and kudoers. More than double the number of employees' posts containing emoticons contain smilies, kudoers use almost the same amount of

Table 5. Most frequently used signals of affect via emoticons and smilies

	Positive				Negative				Neutral			
Rank	E+	*N*	S+	*N*	E-	*N*	S-	*N*	E?	*N*	S?	*N*
1	:)	2,692		3620	!!!	3,272		1057	()	30		851
2	:-)	1,157		2336	???	2,410		463	$$$	21		11
3	;)	779		1462	:(	597		428	(=	18		10
4	:D	588		421	!?!?	252		40	$$	13		10
5	:P	223		37	:-(	123		25	(D)	11		0
6	;-)	208		31	:/	48		18	<=	5		0
7	=)	186		27	;-(	6		18	:>)	4		0
8	8)	102		21	>;	3		18	(8x)	1		0
9	:-D	99		18	:-((	3		11	:-<	1	-	-
10	=>	98		17	:X	2		5	|I	1	-	-

emoticons as smilies. The ratio of usage of emoticons to smilies by gurus and non-kudoer-posts' authors is 1.12 and 1.25 respectively. In not-ranked and non-kudoers' posts this ratio raises to 2.6. Another observation is the disjoint use of emoticons and smilies in posts, the combined count of posts using either emoticons or smilies or both almost equates to the addition of the individual counts of emoticons and smilies in all the groups (%C-(%E+%S) $\approx$ 0).

As non-employees, gurus, ranked and not-ranked users can be less formal than users belonging to the company; it is the case that employees use the forum to make formal announcements where the use of emoticons and smilies would be less obviously appropriate than in more engaging posts. This observation is supported by the fact that employees use less signals of emotion in their posts when compared to the users from the other two role-based groups. Reading Table 6 vertically shows that not-ranked, non-kudoers users and non-kudoer-post authors are the ones who use emoticons with more frequency than their counterparts in all the cases of emoticons; kudoers, ranked users and kudoer-posts' authors are the ones who more use smilies and combined signals of emotion in their respective categorisations.

We show the average frequency and standard deviation of signals of emotion per post for the three categorizations in Table 7, and token average per post containing at least one signal of emotion in the last two columns.

The ratio of frequency of signals of emotion type {positive(+), negative(-), neutral(?)} to overall frequency of signals of emotion per group is shown in Table 8. More than half of the emoticons used by ranked and not-ranked users were negative, although the difference between ratio of positive and negative emoticons for ranked is not as big as for not-ranked users. There is no significant difference between the use of positive and negative emoticons by ranked users in comparison to other categories (p = .3039). Employees and gurus used a significant amount of positive emoticons (more than 80%) compared to negative emoticons. Their use of emoticons is not significantly different (p = .0887 for

Table 6. Number of posts with and without (E)moticons, (S)milies and (C)ombined, with percentage of posts containing these signals of emotion for three categorisations

	Categories	E> 0	No E	%E> 0	S> 0	S= 0	%S> 0	C> 0	C= 0	%C> 0
By role	employee	244	25,246	0.96	572	24,918	2.24	810	24,680	3.18
	guru	1,170	26,319	4.26	1,041	26,448	3.79	2,176	25,313	7.92
	notranked	3,214	47,242	6.37	1,196	49,260	2.37	4310	46,146	8.54
	ranked	5,795	99,054	5.53	5,896	98,953	5.62	11,334	93,515	10.81
User Cat.	kudoer	5,993	133,171	4.31	7,018	132,146	5.04	12,670	126,494	9.10
	nonkudoer	4,430	64,690	6.41	1,687	67,433	2.44	5,960	63,160	8.62
Post Cat.	kudoer	811	17,729	4.37	1,025	17,515	5.53	1,778	16,762	9.59
	nonkudoer	9,612	180,132	5.07	7,680	182,064	4.05	16,852	172,892	8.88

Table 7. Average of signals of emotion per post and tokens per post containing at least one signal of emotion across the three categorisations

		Emoticons		Smilies		Combined		Tokens	
	Category	μ(E)	σ(E)	μ (S)	σ(S)	μ(C)	σ (C)	Mean	sd
By role	employee	1.09	0.351	1.11	0.423	1.112	0.414	78.048	100.481
	guru	1.418	1.076	1.336	0.771	1.402	1	76.622	91.991
	notranked	1.231	0.654	1.144	0.53	1.235	0.673	115.004	143.601
	ranked	1.269	1.619	1.292	1.15	1.321	1.446	92.002	128.743
User Cat.	kudoer	1.273	1.204	1.294	1.094	1.319	1.190	87.88	118.54
	nonkudoer	1.265	1.449	1.143	0.517	1.264	1.303	109.88	145.45
Post Cat.	kudoer	1.29	0.82	1.25	0.98	1.31	0.98	120.83	145.34
	nonkudoer	1.27	1.35	1.27	1.01	1.30	1.25	92.19	125.93

positive and $p = .1879$ for negative). Neutral emoticons are rarely used in this dataset by any user category. Non-kudoers use of negative emoticons is as much as the double of positive, kudoers in the other hand used more positive emoticons but not in the same magnitude as in the role-categorisation. Kudoer posts show more positive emoticon use than non-kudoer posts.

Positive smilies are more used than negative smilies across all groups, with employees using the greatest quantity of positive smilies and not-ranked and non-kudoers, the least. Also employees usage of negative smilies is not significantly different from guru's case ($p = .1129$). Although the use of negative smilies is small in each group, not-ranked and non-kudoer users have the highest ratio of negative smilies, more than double than for gurus. Neutral smilies were rarely used in this dataset.

According to combined signals of emotion ratios, the usage of positive signals reaches more than 56% across employees, gurus, ranked, kudoer users and kudoer and non-kudoer posts, being particularly high in the employees and gurus; their use of positive signals is not significantly different($p = .1271$). Negative signals usage is high in not-ranked and non-kudoer users compared to the other groups. For ranked user there is no significant difference in their use of positive and negative emoticons compared to other categories ($p = .306$). If not stated

Table 8. Ratio of signals of emotion type to total of signals per group

	Category	E(+)	E (-)	E(?)	S(+)	S(-)	S(?)	C(+)	C(-)	C(?)
By role	employee	0.808	0.180	0.011	0.913	0.035	0.052	0.882	0.078	0.040
	guru	0.851	0.147	0.002	0.875	0.052	0.073	0.862	0.104	0.034
	notranked	0.316	0.673	0.011	0.618	0.334	0.048	0.394	0.586	0.020
	ranked	0.480	0.512	0.008	0.709	0.202	0.090	0.596	0.354	0.049
User Cat.	kudoer	0.596	0.396	0.008	0.753	0.161	0.087	0.681	0.268	0.051
	nonkudoer	0.332	0.660	0.009	0.625	0.326	0.050	0.407	0.574	0.019
Post Cat.	kudoer	0.611	0.378	0.011	0.806	0.118	0.076	0.719	0.235	0.046
	nonkudoer	0.473	0.519	0.008	0.720	0.199	0.081	0.583	0.377	0.040

otherwise, all the differences reported with relation to Table 8 are significant ($p < 0.05$).

Table 9 shows the ratio of signals of emotion relativized by the number of posts per group. Usage of positive signals of emotion is high in the group of gurus, followed by ranked and kudoer users. Ratios are very small in the group of employees due to their little usage of emoticons (cf. Table 6). Not-ranked and non-kudoer users have the biggest ratio of negative emoticons compared to positive emoticons (almost double). Positive and negative emoticons are used almost in the same proportion by ranked users. Neutral emoticons usage is very marginal when relativized to number of postings.

Positive smilies are mostly used across all the groups, this is particularly high in guru, ranked, kudoer users and kudoer posts. Neutral emoticons usage is marginal. Combined counts also show a high use of positive signals of emotion than of negative signals of emotion by all groups but not-ranked and non-kudoers where negative signals are mostly used. Nonetheless, the usage of negative emoticons by ranked users is higher than in employees and gurus groups.

Table 9. Ratio of signal of emotion type to total of posts per group

	Category	E(+)	E (-)	E(?)	S(+)	S(-)	S(?)	C(+)	C(-)	C(?)
By role	employee	0.008	0.002	0.000	0.023	0.001	0.001	0.031	0.003	0.001
	guru	0.051	0.009	0.000	0.044	0.003	0.004	0.096	0.011	0.004
	notranked	0.025	0.053	0.001	0.017	0.009	0.001	0.042	0.062	0.002
	ranked	0.034	0.036	0.001	0.051	0.015	0.007	0.085	0.051	0.007
User Cat.	kudoer	0.033	0.022	0.000	0.049	0.010	0.006	0.082	0.032	0.006
	nonkudoer	0.027	0.054	0.001	0.017	0.009	0.001	0.044	0.063	0.002
Post Cat.	kudoer	0.034	0.021	0.001	0.056	0.008	0.005	0.090	0.029	0.006
	nonkudoer	0.030	0.033	0.001	0.037	0.010	0.004	0.067	0.044	0.005

4.2 Correlation with Levels of Interactivity

As seen in Sect. 3.2, one metric of interactivity is the average of postings per individual (APPI). Gurus and kudoers are significantly more involved in posting than employees, not-ranked and non-kudoer users. The APPI values for not-ranked and non-kudoer users are very low, the closest less active group (employees) differs from these values by factor of $\sim$20. Recall that more than 80% of the emoticons and smilies used by gurus are positive. The use of negative signals per not-ranked users is always bigger than their counterparts (e.g. 59.7% in comparison to 8.1 and 12.7% for employees and gurus respectively on combined signals of emotion). Also, non-kuoder users show high use of negative signals (58.6%) compared to kudoer users.

There is little correlation between number of positive emoticons and post length ($r = .05$, $p = .00018$). The correlation of number of negative emoticons and post length was positive but very weak: $r = .28$ ($p < .0001$, two-tailed), while the correlation of number of neutral emoticons and post length was weakly positive ($r = .24$) but significant ($p = .02754$), maybe due to small number of posts that use neutral emoticons ($N = 84$). The correlation of number of positive signals is small ($r = .15$, $p < .0001$). There is a weak positive correlation between number of negative signals of emotion and post length ($r = .30$, $p < .0001$) and for neutral emoticons the correlation is marginal ($r = .15$, $p < .0001$).

A horizontal forum is the one with more new posts than replies to old ones, for instance newsgroups discussing politics (cf. [5]). The forum we explore here is very vertical (more replies than questions) due to its nature: providing support to customers; non-answered threads does not contribute to the image of the forum as a support service mechanism.

5 Concluding Remarks

This exploration of emoticons and smilies has shown some trends according to expertise-oriented user classifications in a technical community forum. Our hypotheses are confirmed with regards to the positive sentiment shown by kudos-receivers. This is directly related to the interactivity variable, since kudos are given as a reward for valuable contributions in the forum. Differences in the usage of two kinds of signals of emotion were found across the different user groups. Negative emoticons are mostly used by common users that often use the forum to expose their technical issues. The findings from this research will be used to explore automatic identification of users deserving promotion.

We plan to explore additional metadata about users and their posts such as: number of views, kudos given by each user, frequency of posting across time and amount of message editing. Another dimension is to explore some intersections between groups, for instance there may be not-ranked users who receive kudos. Also, we have not explored in depth the ranked and not ranked groups as they are assigned ranks according to a hierarchy with promotion based on merit.

As indicated above, we have excluded from this analysis assessment of which posters will eventually be promoted and the extent to which use of signals of affect provide reliable predictors of ultimate promotion.

Acknowledgements. This research is supported by the Trinity College Research Scholarship Program and Science Foundation Ireland (Grant 07/CE/I1142) as part of the Centre for Next Generation Localisation (www.cngl.ie) at Trinity College Dublin. We thank Mikhail Timofeev for providing us with the forum metadata and information about user categories.

References

1. Bifet, A., Holmes, G., Pfahringer, B.: Moa-tweetreader: real-time analysis in twitter streaming data. In: Elomaa, T., Hollmén, J., Mannila, H. (eds.) DS 2011. LNCS, vol. 6926, pp. 46–60. Springer, Heidelberg (2011), http://dl.acm.org/citation.cfm?id=2050236.2050243
2. Carstairs-McCarthy, A.: Synonymy avoidance, phonology and the origin of syntax. In: Hurford, J.R., Studdert-Kennedy, M., Knight, C. (eds.) Approaches to the Evolution of Language. Social and Cognitive Bases, pp. 279–296. Cambridge University Press (1998)
3. Cheng, R., Vassileva, J.: User motivation and persuasion strategy for peer-to-peer communities. In: Proceedings of the 38th Annual Hawaii International Conference on System Sciences, HICSS 2005, vol. 07, p. 193. IEEE Computer Society, Washington, DC (2005), doi:10.1109/HICSS.2005.653
4. Goffman, E.: The Presentation of Self in Everyday Life. Doubleday, New York (1956)
5. Janssen, J., Vogel, C.: Politics makes the Swedish:-) and the Italians:-(. In: Proceedings of LREC – EMOT 2008, Sentiment Analysis: Emotion, Metaphor, Ontology and Terminology, pp. 53–61 (2008)
6. Kucuktunc, O., Cambazoglu, B.B., Weber, I., Ferhatosmanoglu, H.: A large-scale sentiment analysis for Yahoo! answers. In: Proceedings of the Fifth ACM International Conference on Web Search and Data Mining, WSDM 2012, pp. 633–642. ACM, New York (2012), doi:10.1145/2124295.2124371
7. Read, J.: Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In: Proceedings of the ACL Student Research Workshop, ACLstudent 2005, pp. 43–48. Association for Computational Linguistics, Stroudsburg (2005), http://dl.acm.org/citation.cfm?id=1628960.1628969
8. Vogel, C., Janssen, J.F.: Emoticonsciousness. In: Esposito, A., Hussain, A., Marinaro, M., Martone, R. (eds.) COST Action 2102. LNCS, vol. 5398, pp. 271–287. Springer, Heidelberg (2009)