Effect of affective profile on communication patterns and affective expressions in interactions with a dialog system

Marcin Skowron[1], Mathias Theunis[2], Stefan Rank[1], and Anna Borowiec[3]

[1] Austrian Research Institute for Artificial Intelligence, Vienna, Austria
marcin.skowron @ ofai.at, stefan.rank @ ofai.at,
[2] School of Humanities and Social Sciences, Jacobs University, Bremen, Germany
m.theunis @ jacobs-university.de
[3] Gemius SA, Warsaw, Poland
anna.borowiec @ gemius.pl

Abstract. Interlocutors' affective profile and character traits play an important role in interactions. In the presented study, we apply a dialog system to investigate the effects of the affective profile on user-system communication patterns and users' expressions of affective states. We describe the data-set acquired from experiments with the affective dialog system, the tools used for its annotation and findings regarding the effect of affective profile on participants' communication style and affective expressions.

Keywords: affective profile, dialog system, affective computing, HCI

1 Introduction

Emotionally driven online behavior is traceable in a wide range of human communication processes on the Internet. Here, the sum of individual emotions of a large number of users, with their interconnectivity and complex dynamics, influence the formation, evolution and breaking-up of online communities. Our research concentrates on dyadic communication as a fundamental building block for the modeling of more complex, multi-agent communication processes. Using artificial conversational entities, i.e. affective dialog systems, we investigate *the role of emotions* in online, real-time, natural-language-based communication.

In our current research we develop dialog systems and apply them to communicate with members of various e-communities to probe for affective states and background knowledge related to those states (Affect Listeners). These systems communicate with users in a predominantly textual modality, rely on integrated affective components for detecting textual expressions of the users' affective states, and use the acquired information to aid selection and generation of responses. Affect Listeners interact with users via a range of communication channels and interfaces (e.g., Internet Relay Chat (IRC), Jabber, online chat-site interface) and were already integrated as dialog management backbone of a virtual human, the "Virtual Bartender", in a 3D environment [7]. Evaluation results showed that the system ratings for the dialog realism, participants' feeling of an emotional connection with an artificial conversational partner and of chatting enjoyment did not differ from these obtained in a Wizard-of-Oz (WOZ) setting [21].

This paper presents analysis of users communication recorded during new experiments with a revision of the dialog system, based on the "Affect Listeners" platform [19], in a setting typical for online, real-time, text-based communication (i.e., chat-rooms), equipped with three distinct affective profiles. *Artificial affective profile* is defined as a coarse-grained simulation of a personality, corresponding to dominant, extroverted character traits, that can be consistently demonstrated by a system during the course of its interactions with users. In this round of experiments, three distinct affective profiles were provided to the dialog system: positive, negative and neutral. Each affective profile aimed at a consistent demonstration of character traits of the "Affect Bartender" system that could be described as:

- cooperative, emphatic, supporting, positively enhancing, focusing on similarities with a user,
- conflicting, confronting, focusing on differences,
- professional, focused on job, not responding to affective expressions.

Findings related to the effect of affective profiles on the evaluation of the system and self-reported emotional changes experienced during the interaction are presented in[22]. In this paper, we focus on the effect of affective profiles on interaction patterns and participants' expressions of affective states. We consider a set of parameters such as: timing, textual expressions of affective states (as detected by Affective Norms for English Words dictionary (ANEW)[1], Lexicon Based Sentiment Classifier[16], Linguistic Inquiry and Word Count dictionary[17]), dialog act classes and surface features of the participants' communication style (e.g., wordiness, usage of emoticons).

2 Relevant Research

Prior study on the relationship between affective states and dialog patterns observed in the interactions with Intelligent Tutoring Systems, e.g. AutoTutor, was presented in [4]. The study focused on discovering the links between learning and emotions. It applied an emote-aloud procedure in which experiment participants verbalise their affective states experienced in the interaction with the tutoring system. The experimental results demonstrated significant correlations between accuracy of participants' answers and particular affective states, e.g. "confusion" indicating inaccurate answers, "eureka" as an indicator of students learning the material and "frustration" positively correlated with system's negative feedback and negatively correlated with a positive feedback. In our work, a different interaction setting is used, an online virtual bar. Communication content combines task-oriented dialogs specific to the interaction scenario and open domain dialogs regarding participants' attitude and affective responses to current issues of public debate, as well as their affective states expressed during interaction with the system. A further difference to an emote-aloud method: the presented analysis is based on an automated processing and annotation of the acquired dialog logs. In [2], models for utterance selection based on impoliteness that considers emotions, personality and social relations are presented. [13] describes a highly configurable system that generates utterances along the extroversion dimension and reports positive results regarding evaluation.

Taking into consideration limitations of the currently used experimental settings and the applied procedure, the motivation for our work is closer to the goals and visions presented recently e.g., by Picard [18] and Wilks [25]. In particular, the former postulates a change of the focus from comparing average statistics and self-report data across people experiencing emotions in labs to characterising patterns of data from individuals and clusters of similar individuals experiencing emotions in real life. The later stresses the importance of models of human-computer relationship forming a base for long-term interactions that should not be *inherently* tasks-based, e.g., lack of a stopping point to system conversation, role of politeness and users' preferences related to the a specific personality of a system, or its consistency in the long-term relationship.

3 Experimental Settings

3.1 Overview of Dialog System Architecture

The system architecture includes 3 main layers: perception, control and communication. The perception layer provides annotations for user utterances and system response candidates. It applies a set of natural language processing and affective processing tools and resources [20]. Based on the information cues provided by the perception layer, the control layer selects and, if necessary, modifies system response candidates. Further, the layer manages dialog progression taking into account the dialog context and the selected system's affective profile. The control layer uses an information state based dialog management component: Affect Listener Dialog Scripting (ALDS) [19] for the closed-domain and task-oriented parts of the dialog. For the open-domain chats, a template based mechanism and response generation instructions, Affect Bartender AIML set (AB-AIML) [20], are applied. The system's affective profile influences the selection of both ALDS scenarios and subsets of AB-AIML response instructions. To the remaining system response candidates for which no specific affective profile dependent interaction scenarios or system response instructions are provided, an automatic post-precessing is applied, i.e., addition, removal of positive or negative words. The mechanism aims at aligning the affective load, i.e. valence of system response candidates with the selected affective profile [22].

3.2 Characteristic of the participants

The aim of the experiment was to study how affective profiles influence the perception of the system, communication patterns and participants' expressions of affective states. For this purpose invitations were sent to the panelists of a research panel[4]. This is a group of predominantly Polish users who expressed willingness to participate in various on-line surveys. During the study, respondents are in their "natural environment" - a place where they usually use the Internet, which is assumed to make them more receptive as well as spontaneous. For the majority of participants, English, the language in which the experiments were conducted, was not their native language, but all participants who completed the set of interactions had at least average communicative skills in this language. The usage of non-native languages in online interaction environments is a frequent phenomenon and provides the motivation for studying this type of communication. When filling out the registration form, an Internet user provides her demographic data (such as age, gender, education etc.).

Almost 70% of participants who completed the experiment are aged between 24 and 31 (inclusive) and over 90% of participants who completed the experiment are aged between 24 and 39 (inclusive). Over 95% of participants that completed the experiment access the Internet daily or almost daily. Over 70% of them are learning or studying.

3.3 Experimental Procedure

To avoid differences in the evaluation of systems related to the ordering of presentation of the different experimental conditions, the sequence of conditions was randomly and evenly assigned, and the list of evaluation statements was displayed to users before the start of the first interaction so that they could familiarize themselves with the statements to be rated. These statements were related to the following aspects of a completed interaction: chatting enjoyment, feeling of an "emotional connection" with the conversational partner, dialog realism and coherence. Further, participants were asked to report on emotional changes experienced during interaction (i.e., positive, negative) and willingness to chat again with the same partner. During the experiments, after each experimental condition corresponding to an affective profile, participants were asked to express their agreement or disagreement with each of the abovementioned aspects, using a five-point Likert scale from 'strongly disagree' to 'strongly agree'.

Participants completed experiments in an unsupervised manner and were aware that they talk with an artificial dialog system. Interactions were always initiated by an utterance from the system and stopped after 7 minutes, with a suitable closing response followed by the display of the questionnaire. No artificial delays (e.g., a simulation of thinking or typing) were used. System-user interactions were conducted with a web browser based communication interface, similar to popular online chat-rooms, implemented using Javascript and XML-RPC backends (AJAX).

[4] http://www.opinie.pl/

4 Analysed Data-set

Each participant performed three, seven minutes long interactions in a randomized order with three versions of the AffectBartender, introduced above. 91 participants (33 female, 58 male), age between 18 and 42, completed interactions in all three experimental settings resulting in 273 interaction logs.

4.1 Applied Annotation Tools and Resources

The analysis of the presented data-set was conducted with a set of natural language processing and affective processing tools and resources, including: Support Vector Machine Based Dialog Act classifier, Lexicon Based Sentiment Classifier[16], Linguistic Inquiry and Word Count dictionary[17], ANEW dictionary based classifier [1]. Further, we analyzed timing information and surface features of participants communication style such as wordiness and usage of emoticons.

Dialog Act classifier. Dialog act classes are based on the annotation schema used in the NPS Chat Corpus [6]. The originally used taxonomy of DA classes (Accept, Bye, Clarify, Continuer, Emotion, Emphasis, Greet, No Answer, Other, Reject, Statement, Wh-Question, Yes Answer, Yes/No Question), was extended with an additional class "Order" (i.e. for ordering drinks). For this additional class 339 training instances were provided. The original NPS Chat class "System", irrelevant for the system-user dialogs, was excluded along with the set of corresponding training instances. For the presented taxonomy and training set, the Support Vector Machine Based DA classifier achieved 10-fold cross validation accuracy of 76.1%, improving the previously reported classification accuracy for the same data-set achieved with a Maximum Entropy based classifier - 71.2%[20].

Linguistic Inquiry and Word Count - LIWC. This lexical resource provides a classification of words along 64 linguistic, cognitive, and affective categories [17]. Among others, the resource provides 32 word categories for psychological processes (e.g., affective such as positive and negative emotions; cognitive such as insight and causation), 22 linguistic categories (e.g., adverbs, negations, swear words), 7 personal concern categories (e.g., home, work, leisure) 3 paralinguistic dimensions (fillers, assents), for almost 4500 words and word stems. For example, the word "compassion" is categorised in 3 categories: affective processes, positive emotion and social processes; the word "grief" in 4 categories: affective processes, negative emotion, sadness (psychological processes) and death (personal concern). In recent years, LIWC has been successfully applied in various psychological and psycholinguistic studies that included e.g., the investigation of linguistic style, the relations between language use and speaker personality [3].

Sentiment Classifier. Lexicon Based Sentiment classifier[16] provides information on: sentiment class (SC) i.e., negative {-1}, neutral {0}, and positive {1}. Further it assigns positive sentiment value (PS) {+1,...,+5} and negative sentiment value (NS) {-5, ..., -1} to user utterances and system response candidates. The initial scores for the input words are derived from two different emotional word-lists: The "General Inquirer" and "Linguistic Inquiry and Word

Count" (LIWC) dictionary[5], the latter as enriched by [24]. The applied algorithm relies also on a detection of a range of linguistic features such as negation, capitalisation, intensifier, diminisher, etc., which modify the final sentiment score assigned to an input string. Higher absolute values indicate higher emotional content in that dimension and {-1,+1} indicate lack of emotion.

ANEW. Affective Norms for English Words dictionary is based on the assumption that emotion can be defined as a coincidence of values on a number of strategic dimensions [1]. It includes a set of 1,034 commonly used words, including verbs, nouns and adjectives. It provides information on emotional content of an input string, in three affective dimensions: valence, arousal and dominance, on the scale from 1 (very unpleasant, low arousal, low dominance/control) to 9 (very pleasant, high arousal, high dominance/control). For example "abuse" has the following mean score for the three presented affective dimensions (valence - 1.80, arousal - 6,83, dominance - 3.69).

4.2 Effects of Affective Profile on Users' Communication Style

Words and Timing. Repeated measures analyses of variance (ANOVAs) revealed an absence of effect of the affective profile on the number of utterances, words, and characters produced (all Fs(2, 180) < 1.78, ps $> .17$). Restriction in the duration of the interaction between participant and dialog system, as well as the constant responsiveness of the system across conditions most likely explain this result. Participants emitted a mean of 16 utterances ($SD = 7.4$) containing 5 words on average ($SD = 4.3$) during each interaction with the system. Furthermore, the affective profile neither had an effect on the response time (per utterance), $F(2, 180) = .62$, $p = .54$. Participants were equally fast in replying to system's utterances across all conditions.

Dialog Act and LIWC Spoken Categories classes. Omnibus repeated measures ANOVAs showed main effects of the dialog system affective profile on five Dialog Act classes: Statement, Emotion, ynQuestion, Continuer, and yAnswer (Fs(2, 180) > 3.98, ps $< .05$). Pairwise comparisons with Bonferroni correction (see Figure 1) show the expected presence of emotion in the positive compared to the neutral interaction, though the difference is not significant between the neutral an the negative condition [6]. Additionally, the positive profile elicited more statements, less polar questions and less continuations ("and" + text) compared to the negative profile. This higher number of statements and lower number of closed questions might indicate a more successful interaction (i.e., where the user tells more about him/herself and questions the system less), whereas the decrease in continuers remains open to interpretation. Through a similar analysis, a main effect of the affective profile on LIWC Assent class was found, $F(2, 180) = 8.39$, $p < .001$. Specifically, during interactions with the

[5] http://www.liwc.net

[6] In all figures data are normalized with the number of utterances emitted by a user in a given interaction. Asterisks indicate significant differences at $p < .05$. Error bars indicate 1 standard error above and below the mean.

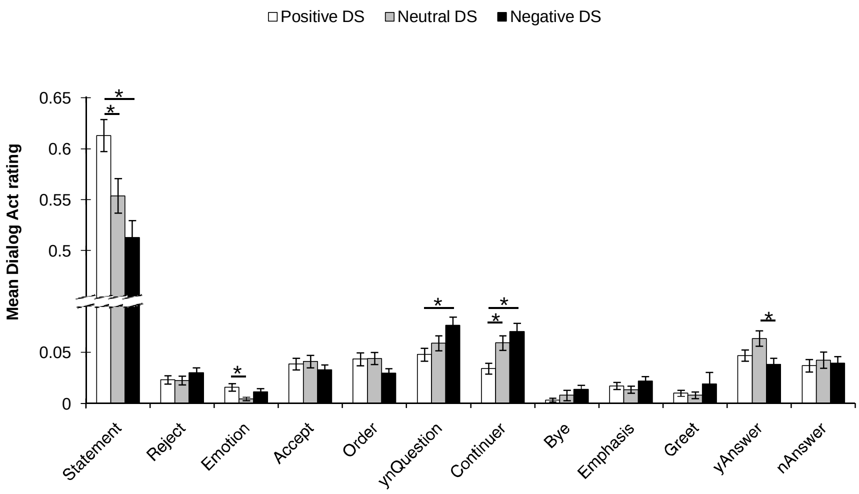

Fig. 1. Mean proportion of Dialog Act classes present in participant's utterances per condition. The Y-axis is broken due to the higher proportion of statements across all utterances, compared to all other classes. DS = Dialog System.

negative profile participants agreed (e.g., "ok", "yes", "yep") significantly less, compared with interactions with the two other profiles. No other significant effect was found on LIWC Spoken Categories classes (Fs(2, 180) $<$ .85, ps $>$ .43).

4.3 Effect of Affective Profile on Users' Expression of Affective States.

After looking at formal aspects of speech, changes in users' affective states were examined through their utterances. [22] showed that user report significant affective changes after each interaction with an artificial affective profile. Investigations were therefore made upon subtle cues of influences of the affective profile on participant's emotions, exploiting the text produced. Based on previous research on emotional contagion [9], it was hypothesized that the dialog system's valence would linearly affect user's emotional state. In other terms, we expected to find changes toward a more negative emotional state in the user, elicited by exchanges with the dialog system's negative profile. The reverse effect was expected to be found in exchanges with the positive profile, and an absence of change was predicted for interactions with the neutral profile.

Emoticons, Sentiment Classifier, and ANEW lexicon. A first confirmation of the abovementioned hypothesis was found in a significant effect of the affective profile on user's production of positive emoticons, $F(2, 180) = 9.02$, $p <$ .001. Pairwise comparisons reveal that users emitted significantly more positive emoticons while interacting with the positive affective profile, compared with interactions with the two other profiles. No effect was found concerning negative emoticons production, $F(2,\ 180) = 2.41$, $p = .09$. Furthermore, it was found that the dialog profile significantly affects the positive Sentiment Value found in users' utterances, $F(2, 180) = 15.08$, $p < .001$. As depicted in Figure 2 (Panel A), participants interacting with the negative profile produced text classified as significantly less positive compared with the two other conditions. No significant effect was found concerning the negative Sentiment Value, $F(2, 180) = .64$, $p = .53$. Additionally, the affective profile of the system was found to have a significant impact on valence, arousal, and dominance of user's utterances, based on ANEW ratings (Fs(2, 180) > 19.23, ps $< .001$). Compared with the two other conditions, when communicating with the negative profile, participants emitted utterances classified as significantly less positive, less activated, and less dominant (see Figure 2, Panel B).

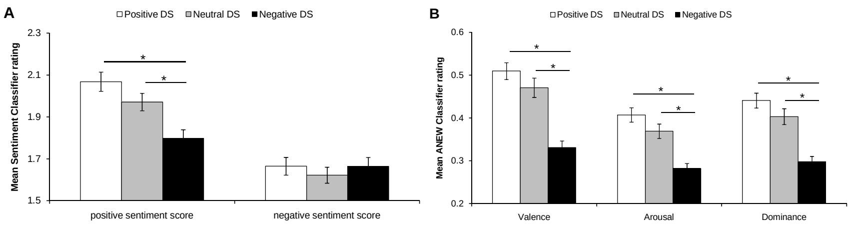

Fig. 2. Valence, arousal, and dominance ratings found in user exchanges with the dialog system (DS). Panel A shows the mean positive and negative Sentiment Classifier score per condition. Panel B shows the mean valence, arousal, and dominance scores based on the ANEW lexicon.

LIWC Psychosocial Processes classes. Finally, user's utterances were also analyzed using LIWC Psychosocial Processes classes. As hypothesized, the affective profile was found to have several significant effects on Affective Processes detected in text with LIWC's lexicon (Fs(2, 180) > 3.56, ps $< .05$). Multiple pairwise comparisons showed—among others—that, during interactions with the negative profile, users used significantly less positive emotion words (e.g., "love", "nice", "sweet"), more negative emotion words (e.g., "ugly", "nasty", "sad"), and more anger-related words (always compared with interactions with the two other profiles, positive and neutral).

5 Discussion and Outlook

Creating systems which adequately detect and respond to human emotions can have profound consequences. As research on emotional contagion showed [9], individuals tend to synchronize their affective states. When one interacts with a happy person, there is a higher probability that one will get happy as well [8], compared to the probability of getting upset or afraid. Social network analysis demonstrated that emotions not only spread from person to person, but throughout entire networks [10]. Moreover, it has now been clearly demonstrated that written text is affected by emotional states [23]. Research showed that emotional valence can accurately be detected in text [24], and linked to an individual's affective state [11]. Recent developments even evidenced the possibility to detect emotional categories from text (e.g., fear, anger, shame)[15]. Taking into account both streams of research, one showing that emotion is contagious, the other that it can be accurately detected, our approach has a high potential for enriching user's experience, and beyond. Combining an accurate detection of emotion in user text with an adequate emotional response from the system can enrich communication at a point close to human-human interaction [14]. Looking at the effect of such a design, the above described study presents an attempt at grasping the far-reaching implications of the development of an emotionally intelligent system. Effort now has to be put into increasing the level of compliance of the system architecture with psychological research [5], as well as taking into account the complex variations in user's affective experience [12]. Our future research includes further investigation of the effect of emotions in user-system interactions, e.g., social sharing of emotion, self-disclosure, both in the single and multiple users interaction environments.

Acknowledgments

The work reported in this paper is partially supported by the European Commission under grant agreements CyberEmotions (FP7-ICT-231323) and IRIS (FP7-ICT-231824). The Austrian Research Institute for AI is supported by the Austrian Federal Ministry for Transport, Innovation, and Technology.

References

1. M.M. Bradley and P.J. Lang. Affective norms for english words (anew): Stimuli, instruction manual and affective ratings. Univ. of Florida, 1999.
2. S. Campano and N. Sabouret. A socio-emotional model of impoliteness for non-player characters. In *Proc. of the 8th Int. Conf. on Autonomous Agents and Multiagent Systems*, AAMAS09, pages 1123–1124, Richland SC, 2009. Int. Foundation for Autonomous Agents and Multiagent Systems.
3. C. K. Chung and J. W. Pennebaker. Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *J.of Research in Personality*, 42:96–132, 2008.
4. S. D'Mello, S. Craig, A. Witherspoon, J. Sullins, B. McDaniel, B. Gholson, and A. Graesser. The relationship between affective states and dialog patterns during interactions with autotutor. In Griff Richards, editor, *Proc. of World Conf. on E-Learning in Corporate, Government, Healthcare, and Higher Education 2005*, pages 2004–2011. AACE, October 2005.
5. J.R.J. Fontaine, K.R. Scherer, E.B. Roesch, and P.C. Ellsworth. The world of emotions is not two-dimensional. *Psychological Science*, 18(12):1050–1057, 2007.
6. E. Forsyth and C. Martell. Lexical and discourse analysis of online chat dialog. In *Proc. of the First IEEE Int. Conf. on Semantic Computing*, pages 19–26, 2007.

7. S. Gobron, J. Ahn, S. Quentin, D. Thalmann, M. Skowron, S. Rank, G. Paltoglou, M. Thelwall, and A. Kappas. 3d-emochatting: an interdisciplinary communication model for vr chatting. *in review*, submitted.
8. E. Hatfield, J.T. Cacioppo, and R.L. Rapson. Emotional contagion. *Current Directions in Psychological Science*, 2(3):96–99, 1993.
9. E. Hatfield, J.T. Cacioppo, and R.L. Rapson. *Emotional Contagion.* Cambridge Univ. Press, 1994.
10. A.L. Hill, D.G. Rand, M.A. Nowak, and N.A. Christakis. Emotions as infectious diseases in a large social network: the sisa model. *Proc. of the Royal Society B*, 277(1701):3827–3835, 2010.
11. A. Kappas, D. Kuester, M. Theunis, and E. Tsankova. Cyberemotions: Subjective and physiological responses to reading online discussion forums. In *Society for Psychophysiological Research Abstracts for the Fiftieth Annual Meeting*, 2010.
12. P. Kuppens, Z., and F. Tuerlinckx. Feelings change: Accounting for individual differences in the temporal dynamics of affect. *Journal of Personality and Social Psychology*, 99(6):1042–1060, 2010.
13. F. Mairesse, M. Walker, M. Mehl, and R. Moore. Using linguistic cues for the automatic recognition of personality in conversation and text. *J.of Artificial Intelligence Research*, 30:457–500, 2007.
14. Albert Mehrabian and James A. Russell. *An Approach to Environmental Psychology.* MIT Press, 1974.
15. A. Neviarouskaya, H. Prendinger, and M. Ishizuka. Affect analysis model: novel rule-based approach to affect sensing from text. *Natural Language Engineering*, 17(1):95–135, 2011.
16. G. Paltoglou, S. Gobron, M. Skowron, M. Thelwall, and D. Thalmann. Sentiment analysis of informal textual communication in cyberspace. In *In Proc. Engage 2010, Springer LNCS State-of-the-Art Survey*, pages 13–25, 2010.
17. J. W. Pennebaker, M. E. Francis, and R. K. Booth. *Linguistic Inquiry and Word Count: LIWC 2001.* Erlbaum Publishers, 2001.
18. R.W. Picard. Emotion research by the people, for the people. *Emotion Review*, 2, 2010.
19. M. Skowron. Affect listeners. acquisition of affective states by means of conversational systems. In *Development of Multimodal Interfaces - Active Listening and Synchrony*, Lecture Notes in Computer Science, pages 169–181. Springer, 2010.
20. M. Skowron and G. Paltoglou. Affect bartender - affective cues and their application in a conversational agent. In *IEEE Symposium Series on Computational Intelligence 2011, Workshop on Affective Computational Intelligence.* IEEE, 2011.
21. M. Skowron, H. Pirker, S. Rank, G. Paltoglou, J. Ahn, and S. Gobron. No peanuts! affective cues for the virtual bartender. In *Proc. of the Florida Artificial Intelligence Research Society Conf.* AAAI Press, 2011.
22. M. Skowron, S. Rank, M. Theunis, and J. Sienkiewicz. The good, the bad and the neutral: affective profile in dialog system-user communication. In *in a review.*
23. Yla R. Tausczik and James W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *J.of Language and Social Psychology*, 29(1):24–54, 2010.
24. M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A Kappas. Sentiment strength detection in short informal text. *J.of the American Society for Information Science and Technology*, 61(12):2544–2558, 2010.
25. Yorick Wilks. Is a companion a distinctive kind of relationship with a machine? In *Proc. of the 2010 Workshop on Companionable Dialogue Systems*, pages 13–18, Uppsala, Sweden, July 2010. Association for Computational Linguistics.