

Automatic identification of writers' intentions: Comparing different methods for predicting relationship goals in online dating profile texts

Chris van der Lee, Tess van der Zanden, Emiel Krahmer, Maria Mos, and Alexander Schouten

Tilburg center for Cognition and Communication (TiCC),

School of Humanities and Digital Sciences, Tilburg University

{c.vdrlee, t.vdrzanden}@tilburguniversity.edu

{maria.mos, a.p.schouten, e.j.krahmer}@tilburguniversity.edu

Abstract

Psychologically motivated, lexicon-based text analysis methods such as LIWC (Pennebaker et al., 2015) have been criticized by computational linguists for their lack of adaptability, but they have not often been systematically compared with either human evaluations or machine learning approaches. The goal of the current study was to assess the effectiveness and predictive ability of LIWC on a relationship goal classification task. In this paper, we compared the outcomes of (1) LIWC, (2) machine learning, and (3) a human baseline. A newly collected corpus of online dating profile texts (a genre not explored before in the ACL anthology) was used, accompanied by the profile writers' self-selected relationship goal (long-term versus date). These three approaches were tested by comparing their performance on identifying both the intended relationship goal and content-related text labels. Results show that LIWC and machine learning models both correlate with humans in terms of content-related label assignment. Furthermore, LIWC's content-related labels corresponded more strongly to humans than those of the machine learning model. Moreover, all approaches were similarly accurate in predicting the relationship goal.

1 Introduction

When investigating large textual datasets, it is oftentimes necessary to use (automated) tools in order to make sense of the texts. Such tools can help expose properties of texts or of the texts' author (Riffe et al., 2014). A distinction in these tools can be made between predefined lexicon-based approaches and more content-specific machine learning approaches. One commonly used lexicon-based approach is the Linguistic Inquiry and Word Count program (LIWC; Pennebaker et al., 2015). This approach assigns words to

one or more (psychologically validated) labels associated with the word. These labels might reveal more about a writer's thought processes, emotional states, and intentions (Tausczik and Pennebaker, 2010).

Text analysis tools such as LIWC have become more popular with the surge of social media: researchers want to assess, for instance, the sentiment of social media users on various matters, and lexicon-based text analysis tools can provide help with that. At the same time, these tools have also garnered criticism, for example, because they do not differentiate between domains and cannot deal with non-literal language use (e.g., irony), or out-of-vocabulary terms frequently seen within noisy text (e.g., typos or (internet) slang) (Panger, 2016; Franklin, 2015; Schwartz et al., 2013). This is something that machine learning methods might be better suited for as they can be trained on specific content, thus are able to analyze more complex language. Yet, not much is known about the effectiveness of lexicon-based compared to machine learning methods or a ground truth: comparative research is scarce, with few exceptions like Hartmann et al. (2019). Thus, outcomes of lexicon-based approaches are often taken at face value, without knowing how they compare to human attributions. While some researchers dispute the effectiveness of lexicon-based approaches (Kross et al., 2019; Johnson and Goldwasser, 2018), there are others who found that such approaches are helpful on their own (Do and Choi, 2015), or that classification performance increases with the addition of features from such approaches (Sawhney et al., 2018; Pamungkas and Patti, 2018). Additionally, most work on writer's intentions focuses on basic emotions only (Yang et al., 2018; Chen et al., 2018; Yu et al., 2018). Thus, LIWC's wide range of psychology-related label detection is presently not matched by others.

The social media domain, which the online dating domain (hereafter: dating profiles) is part of, might be challenging for LIWC, since these texts often contain non-standard language and noise. LIWC may nevertheless be a viable tool for analyzing dating profiles. Previous research has found that intended relationship goals are related to psychological traits (Feeney and Noller, 1990; Peter and Valkenburg, 2007), and that dating profiles can contain information about a writer’s psychological and mental states (Ellison et al., 2006). This underlying psychological layer is something that may be exploited by LIWC, since previous research found that the tool can expose such psychological and mental states from linguistic behavior (Tausczik and Pennebaker, 2010; Van der Zanden et al., 2019).

The goal of the current study was to assess the effectiveness and predictive ability of LIWC on a relationship goal classification task. For this, LIWC was compared to human judgment and machine learning approaches in three steps. First, the quality of LIWC’s content-related labels was assessed by comparing the values given to content-related labels to those of humans and a regression model. Second, the meaningfulness of LIWC’s dictionary was investigated by using the label values as features for a classification model that predicts relationship type, contrasting these results with the predictions of humans and a classification model using word features. Third, a qualitative evaluation based on topic models, Gini Importance scores, and log-likelihood ratios was conducted to find limitations of LIWC’s lexicon.

2 Method

2.1 Corpus

Corpus	Long-term			Date		
	Texts	Tokens	Types	Texts	Tokens	Types
Full	10,696	863,227	32,020	1,634	127,644	11,274
Train+val	1,464	117,947	9,973	1,464	115,227	10,540
Test	150	11,886	2,383	150	11,738	2,592

Table 1: Descriptives of the dating profile corpus

A total sample of 12,310 dating profiles together with the indicated desired relationship goal was collected from a popular Dutch dating site (see Table 1). These profiles were anonymized after collection, and were between 50 and 100 words, written in Dutch ($M = 80.36$ words, $SD = 14.56$).

Ethical clearance was obtained from the university for the collection of the dating profiles and the use for further text analysis. The (anonymized) corpus itself and the results from the human evaluation are available upon request.

2.2 LIWC

LIWC 2015 (Pennebaker et al., 2015) was used for the experiments, with the Dutch lexicon by Van Wissen and Boot (2017). This Dutch version of LIWC is of similar size as the English version and the scores have been found to correlate well with those of its English counterpart when tested on parallel corpora (Van Wissen and Boot, 2017). LIWC works by iterating over all words and multi-word phrases in a text and checking whether the word or phrase is in the predefined lexicon of one or more labels. There are 70 labels in total. LIWC outputs percentage scores. For example, if 8% of words are an I-reference, the I-reference score would be 8.

2.3 Human Evaluation

In this study, 152 university students participated (68% female, mean age = 21.8 years). For their voluntary participation they received course credit. Altogether, these participants rated a random sample of 300 profile texts (test set in Table 1). Each participant judged six texts in total: 3 texts where the indicated goal was a long-term relationship, and 3 texts where this goal was a date. Approximately three judgments for each of the 300 profile texts were obtained.

Participants rated the degree to which the profile writer discussed six topics which were deemed important for dating profiles based on previous research (status, physical appearance, positive emotion, I-references, you-references, we-references; see Appendix A) (Davis and Fingerman, 2016; Groom and Pennebaker, 2005; Van der Zanden et al., 2019). These ratings were done on six items, all 7-point Likert scales (“To what degree is the writer of the text talking about: status related qualities (e.g., job, achievements), physical qualities (e.g., height, build), positive emotions, themselves (use of ‘I’), the reader (use of ‘you’), a group the writer belongs to (use of ‘we’)”, ranging from “low degree” to “high degree”). The judgments were used as a baseline for the label assignment task. Furthermore, participants indicated whether they thought the profile text writer sought for a long-term relationship or a date (Krippendorff’s α

Method	Status	Physical appearance	Positive emotion	I	You	We
LIWC	.42**	.20**	.30**	.45**	.44**	.36**
Regression	.51**	.13*	.13*	.31**	.37**	.29**

Table 2: Pearson’s r compared to humans. Vertically aligned bold values differ significantly. * $p < .05$, ** $p < .001$.

= 0.24). These predictions were used as a baseline for the relationship goal classification task. Additionally, participants were asked to highlight the words in the text on which they based their long-term or date prediction. All marked words were then collected and counted for the qualitative analysis.

2.4 Label Assignment Task

The goal of this task was to evaluate the similarity of labels from lexicon-based and machine learning approaches compared to a human baseline. The output of LIWC was limited to the six labels discussed in Section 2.3. The 300 dating profiles evaluated in the human evaluation task were rated by LIWC for fair comparison.

The same 300 texts were also used (with random ten-fold cross-validation) for the regression model. This model was trained to give continuous scores on the six text labels. Word features were chosen for fair comparison, since LIWC is word-based and humans also tend to analyze texts at word, phrase, or sentence level (Marsi and Kraemer, 2005). TheilSenRegressor was the regression algorithm used (see Appendix B for details).

2.5 Relationship Goal Identification Task

With this task, the meaningfulness of the lexicons used by LIWC to capture writers’ relationship goals was investigated and compared to the feature sets that humans and machine learning approaches use. To do so, three classification models were used. One classification model used LIWC’s label scores on the aforementioned six labels as features. The second classification model used word features. Furthermore, a meta-classifier was trained on the probability scores of the classification model with LIWC features and the model with word features. This was done to investigate if LIWC and word features use different facets of a text to distinguish between relationship goals. If so, pooling them together could achieve some kind of synergy, resulting in higher accuracy scores.

Metric	Human	LIWC	Word	Meta
Precision	0.55	0.54	0.61	0.61
Recall	0.67	0.87	0.65	0.65
F1	0.60	0.67	0.63	0.63
Accuracy	0.57	0.56	0.62	0.62

Table 3: Accuracy scores by humans and classification models. Bold indicates highest score on metric.

A total of 3,228 texts (1,614 texts for each relationship group; see Table 1) was used for training and testing. This sample was randomly stratified on gender, age and education level based on the distribution of the group of date seekers. For the classification models, the text was trained using 2,635 texts and validated on 293 texts. Finally, to enable fair comparison between methods, the model was tested on the 300 texts rated by humans. While this test dataset is relatively small, only minor differences were found between accuracy scores when trained on the full dataset using ten-fold cross-validation (approximately 1-2%). Thus, the test set was sufficiently large to obtain relatively stable results. Furthermore, a test was done using Dutch word2vec word-embeddings pre-trained on the COW corpus (Řehůřek and Sojka, 2010; Tulkens et al., 2016) for the classification model with word features, but this did not lead to an increase of accuracy scores. For all classification models, an LSTM network with eight layers was used (see Appendix B for details).

2.6 Qualitative analysis

A qualitative analysis of the output on the relationship goal identification task was performed to analyze possible shortcomings of LIWC’s lexicon. Indicative words for identification according to Gini Importance scores obtained with XGBoost were compared to LIWC’s lexicon (Breiman et al., 1984). Furthermore, LIWC’s lexicon was compared to indicative words according to humans and according to log-likelihood ratio scores (Dunning, 1993). Labels from LIWC were also compared with topics obtained by topic modeling (see Appendix B).

3 Results

3.1 Label Assignment Task

For the label assignment task, the performance of LIWC and the regression model were measured using (two-tailed) Pearson’s r . Results show that

Topic name	Words
Person characteristics	I, humor, good, man, honest, sportive, like, social, sometimes, sweet, spontaneous, woman, positive, little bit, open, stubborn, especially, reliable, human, happy
Self-disclosure	I, you, none, much, say, come, good, something, myself, human, find, become, we, think, sit, always, see, sometime, other
Being together	I, you, want, woman, someone, together, man, sweet, find, know, with, you, good, each other, relationship, spontaneous, cozy, go, where, come
Family and occupation	I, LOCATION, PERSON, reside, year, name, visit, ORGANISATION, work, child, you, come, MISCELLANEOUS, go, time, since, like, old
Named entities, English, and Misc.	<i>pro</i> , MISCELLANEOUS, LOCATION, PERSON, <i>you, the</i> , ORGANIZATION, <i>to, a, none, and, â, what, I, one, life</i> , rather, music, EVENT, <i>my</i>
Nature and hobbies	I, like, find, nice, go, friend, eat, movie, cozy, time, watch, couch, delightful, walk, evening, day, good, make, enjoy, music

Table 4: Translated topic models with assigned topic names.

both LIWC and the regression model correlate significantly with human behavior for all six investigated labels. This suggests that LIWC and a regression model can obtain label scores similar to humans. However, it should be noted that the correlation coefficients are relatively low (ranging from .13 to .51), which indicates a weak to moderate relationship between the regression models and human judgments.

Fishers r to z transformation was employed to investigate whether the strength of the correlation with humans differed significantly between LIWC scores and word-based regression scores. Overall, LIWC performed better on this task: the correlation for LIWC on positive emotions ($p = .03$) as well as I-references ($p = .05$) was significantly stronger than the correlation scores for the regression model on these labels (see Table 2). This indicates that LIWC scores are more similar to the label scores of human annotators (at least for positive emotions and I-references) than the scores of the regression model.

3.2 Relationship Goal Identification Task

For the intended relationship goal identification task, chi-square tests were performed on the predictions for all different methods, to compare them to chance and to each other. All methods turned out to perform better than chance (humans: $\chi^2(1) = 17.58$, $p < .001$; word-based classifier: $\chi^2(1) = 17.28$, $p < .001$; classifier with LIWC features: $\chi^2(1) = 5.33$, $p = .02$; meta-classifier: $\chi^2(1) = 17.28$, $p < .001$). These results suggest that humans, LIWC, and a word-based regression model are similarly accurate in identifying a writer’s relationship goal. This was further corroborated by a 4 (text analysis method) x 2 (correct vs. incorrect judgments) not significant chi-square test ($\chi^2(3) = 4.22$, $p = .24$), meaning that there was no method that performed significantly better than any other

method (see Table 3).

Accuracy for the meta-classifiers did not increase for the relationship goal identification task. The accuracy score of the meta-classifier was the same as the word-based classifier, which suggests that LIWC features and word features pick up on the same aspects of the text. Since the classification model with word features performed slightly better, the meta-classifier likely learns to focus on the probability scores of that model.

3.3 Qualitative Analysis

With all 70 labels, LIWC manages to capture only 15% of the types in the dating profile texts, which suggests that a substantial amount of information is not captured by the approach. Information that is missing are words such as ‘date’, ‘profile’, ‘click’, and ‘friendship’ (all $\chi^2(1) \geq 5.39$, and $p \leq .02$): important relationship-related words, and good discriminators according to the word-based classification model, humans, and log-likelihood ratio. This illustrates that LIWC was not necessarily built with dating profiles in mind.

Distinctive words like ‘spontaneous’ ($\chi^2(1) = 0.13$, $p = .72$, but distinctive according to humans and Gini Importances), ‘caring’, ‘quiet’, ‘honest’, and ‘sweet’ (all $\chi^2(1) \geq 7.18$, and $p \leq .007$) show that LIWC is missing a person characteristics category. Topic modeling also found a personality traits topic, further emphasizing the importance of this label. Similarly, words like ‘sea’ and ‘nature’ (both $\chi^2(1) \geq 5.99$ and $p = .01$), and a nature-related topic model shows that a nature-focused label is important to distinguish relationship goals, which LIWC is currently lacking (see Table 4 and Table 5).

However, while there are some systematic patterns to be found regarding what LIWC is not capturing, do note that LIWC’s scores on the two

tasks were similar to machine learning and to humans. This suggests that the relatively small percentage of word types that LIWC is processing is meaningful. The top 100 most important features according to log likelihood ratios, humans, and Gini Importances corroborates this suggestion. 62% of the top 100 most important words according to log likelihood ratios are found in LIWC, 81% of the top 100 most important words according to humans, and 90% of the top 100 most important words according to Gini Importances.

Method	Important features
Human	seek, date, nice, know, people, undertake, spontaneous, sweet, terrace, pretty, enjoy, child wish, family person
Classifier	date, spontaneous, let, live, nature, sociable, send, build, exercise, nice, independent, again, friendship, sea, girl, terrace
LLR	quiet, sweet, nothing, nature, fetch, again, profile, click, feel free, weekend, sea, people, visit, caring

Table 5: Translated words not in LIWC’s lexicon ordered by importance for relationship goal identification. Blue is indicative of long-term, red is indicative of date

4 Discussion

In this study, a lexicon-based text analysis method (LIWC) was compared to machine learning approaches (regression, classification model), with human judgment scores as a baseline. Lexicon-based methods are criticized because they may not capture complex elements of language and do not discriminate between domains. Still, research often takes the outcomes of these approaches at face value without assessing whether they accurately reflect reality. This study aimed to address these issues using three tasks: (1) assigning content-related labels to texts, (2) predicting intended relationship goals, and (3) comparing the output of the different approaches with a qualitative study. While (1) was used to investigate if LIWC’s labels reflect reality, (2) and (3) aimed to elucidate if LIWC’s labels are sufficient to highlight differences in intended relationship goals. The three tasks were conducted on a newly collected corpus of online dating profiles.

The results of this study show that LIWC is a viable text analysis method for these tasks. Despite the fact that it uses a fixed word list and therefore might miss context and out-of-vocabulary words, it performed similarly to machine learning methods and humans. The label assignment

task showed that the labels of LIWC and the regression model both correlated with the labels assigned by humans. Furthermore, for some labels, LIWC’s scores corresponded more to human judgments than those of the regression model. This suggests that LIWC’s lexicon was chosen meaningfully and that despite its limitations, it seems to be good at exposing textual themes. This is corroborated by the fact that most of the important words according to Gini Importances, log likelihood ratio, and humans were in LIWC’s lexicon.

However, it should be noted that the sample size for this task was small (300 texts), and that results could be different if there was more training data. Relationship goal prediction turned out to be a difficult task (low accuracy scores overall, and low inter-rater agreement). Thus, future research should look into extending the human evaluation dataset with more texts and judgments per text. Nevertheless, humans and all classification models scored similarly on accuracy and performed above chance, suggesting that LIWC does cover categorical differences between long-term relationship and date seekers, although LIWC seems to pick up on the same signal as the word-based classification model. Results from the qualitative analysis show that the categories in LIWC might not be sufficient to cover the full range of categorical linguistic differences between the two groups. These shortcomings might be addressed by novel approaches that aim to combine dictionaries with neural text analysis methods, such as Empath (Fast et al., 2016). Or by extending neural Emotion Classification and Emotion Cause Detection systems like (Yang et al., 2018; Chen et al., 2018; Yu et al., 2018) to cover more psychology-relevant categories. Using novel pre-trained word-embeddings such as BERT (Devlin et al., 2019) could also boost the results for the current approach, as this has improved results for many tasks.

The focus of this study on the intended relationship goals of online daters was a challenge that had not been investigated in previous computational linguistics research. We must note that this is just one example of a task for which LIWC could be used. Studies have shown that LIWC may be less suited for some tasks, such as sentiment analysis (Hartmann et al., 2019). However, the current results indicate that it can be a viable method for tasks that tend to look at other, deeper, psychological constructs.

Acknowledgements

We received support from RAAK-PRO SIA (2014-01-51PRO) and The Netherlands Organization for Scientific Research (NWO 360-89-050), which is gratefully acknowledged. We would also like to thank the anonymous reviewers for their valuable and insightful comments.

References

- Bing Bai. 2017. [LSTM with word2vec embeddings](#).
- L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA.
- Ying Chen, Wenjun Hou, Xiyao Cheng, and Shoushan Li. 2018. Joint learning for emotion classification and emotion cause detection. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 646–651.
- François Chollet et al. 2015. [Keras](#).
- Xin Dang, Hanxiang Peng, Xueqin Wang, and Heping Zhang. 2008. Theil-sen estimators in a multiple linear regression model. *Olemiss Edu*.
- Eden M Davis and Karen L Fingerman. 2016. Digital dating: Online profile content of older and younger adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 71(6):959–967.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. [BERT: Pre-training of deep bidirectional transformers for language understanding](#). In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Hyo Jin Do and Ho-Jin Choi. 2015. Korean Twitter emotion classification using automatically built emotion lexicons and fine-grained features. In *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation: Posters*, pages 142–150.
- Ted Dunning. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74.
- Nicole Ellison, Rebecca Heino, and Jennifer Gibbs. 2006. Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication*, 11(2):415–441.
- Ethan Fast, Binbin Chen, and Michael S Bernstein. 2016. Empath: Understanding topic signals in large-scale text. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 4647–4657. ACM.
- Judith A Feeney and Patricia Noller. 1990. Attachment style as a predictor of adult romantic relationships. *Journal of Personality and Social Psychology*, 58(2):281–291.
- Emma Franklin. 2015. Some theoretical considerations in off-the-shelf text analysis software. In *Proceedings of the Student Research Workshop associated with RANLP 2015*, pages 8–15.
- Carla J Groom and James W Pennebaker. 2005. The language of love: Sex, sexual orientation, and language use in online personal advertisements. *Sex Roles*, 52(7-8):447–461.
- Jochen Hartmann, Juliana Huppertz, Christina Schamp, and Mark Heitmann. 2019. Comparing automated text classification methods. *International Journal of Research in Marketing*, 36(1):20–38.
- Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural Computation*, 9(8):1735–1780.
- Kristen Johnson and Dan Goldwasser. 2018. Classification of moral foundations in microblog political discourse. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 720–730.
- Ethan Kross, Philippe Verduyn, Margaret Boyer, Brittany Drake, Izzy Gainsburg, Brian Vickers, Oscar Ybarra, and John Jonides. 2019. Does counting emotion words on online social networks provide a window into peoples subjective experience of emotion? A case study on Facebook. *Emotion*, 19(1):97–107.
- Erwin Marsi and Emiel Krahmer. 2005. Classification of semantic relations by humans and machines. In *Proceedings of the ACL workshop on Empirical Modeling of Semantic Equivalence and Entailment*, pages 1–6. Association for Computational Linguistics.
- Andrew K McCallum. 2002. [MALLET: A machine learning for language toolkit](#). <http://mallet.cs.umass.edu>.
- Endang Wahyu Pamungkas and Viviana Patti. 2018. Exploiting emojis and affective content for irony detection in English tweets. In *International Workshop on Semantic Evaluation*, pages 649–654. Association for Computational Linguistics.
- Galen Panger. 2016. Reassessing the Facebook experiment: Critical thinking about the validity of Big Data research. *Information, Communication & Society*, 19(8):1108–1126.

- James W Pennebaker, Roger J Booth, and Martha E Francis. 2015. *Linguistic Inquiry and Word Count: LIWC2015*. <http://www.liwc.net> Pennebaker Conglomerates, Austin, TX.
- Jochen Peter and Patti M Valkenburg. 2007. Who looks for casual dates on the internet? A test of the compensation and the recreation hypotheses. *New Media & Society*, 9(3):455–474.
- Radim Řehůřek and Petr Sojka. 2010. **Software Framework for Topic Modelling with Large Corpora**. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta. ELRA.
- Daniel Riffe, Stephen Lacy, and Frederick Fico. 2014. *Analyzing media messages: Using quantitative content analysis in research*. Routledge.
- Ramit Sawhney, Prachi Manchanda, Raj Singh, and Swati Aggarwal. 2018. A computational approach to feature extraction for identification of suicidal ideation in tweets. In *Proceedings of ACL 2018, Student Research Workshop*, pages 91–98.
- Hansen Andrew Schwartz, Johannes Eichstaedt, Eduardo Blanco, Lukasz Dziurzyński, Margaret L Kern, Stephanie Ramones, Martin Seligman, and Lyle Ungar. 2013. Choosing the right words: Characterizing and reducing error of the word count approach. In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics*, volume 1, pages 296–305.
- Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical bayesian optimization of machine learning algorithms. In *Advances in neural information processing systems*, pages 2951–2959.
- Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54.
- Laure Thompson and David Mimno. 2018. Authorless topic models: Biasing models away from known structure. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3903–3914.
- Stephan Tulkens, Chris Emmery, and Walter Daelemans. 2016. Evaluating unsupervised Dutch word embeddings as a linguistic resource. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, Paris, France. European Language Resources Association.
- Leon Van Wissen and Peter Boot. 2017. An electronic translation of the LIWC dictionary into Dutch. In *Proceedings of eLex 2017 conference*, pages 703–715.
- Yang Yang, Zhou Deyu, and Yulan He. 2018. An interpretable neural network with topical information for relevant emotion ranking. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3423–3432.
- Jianfei Yu, Luis Marujo, Jing Jiang, Pradeep Karuri, and William Brendel. 2018. Improving multi-label emotion classification via sentiment classification with dual attention transfer network. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1097–1102.
- Tess Van der Zanden, Alexander P Schouten, Maria Mos, Chris Van der Lee, and Emiel Kraahmer. 2019. Effects of relationship goal on linguistic behavior in online dating profiles: A multi-method approach. *Frontiers in Communication*, pages 1–30.