The Role of Emotions in Native Language Identification

Ilia Markov[1], Vivi Nastase[2], Carlo Strapparava[3], Grigori Sidorov[4]

[1]INRIA, Paris, France
[2]University of Heidelberg, Heidelberg, Germany
[3]Fondazione Bruno Kessler, Trento, Italy
[4]Instituto Politécnico Nacional, Center for Computing Research, Mexico City, Mexico
`ilia.markov@inria.fr, nastase@cl.uni-heidelberg.de,`
`strappa@fbk.eu, sidorov@cic.ipn.mx`

Abstract

We explore the hypothesis that emotion is one of the dimensions of language that surfaces from the native language into a second language. To check the role of emotions in native language identification (NLI), we model emotion information through polarity and emotion load features, and use document representations using these features to classify the native language of the author. The results indicate that emotion is relevant for NLI, even for high proficiency levels and across topics.

1 Introduction

Native Language Identification (NLI) is the task of identifying the native language (L1) of a person based on his/her writing in the second language (L2). NLI can inform security, marketing and educational applications by tuning pedagogical materials to L1s, and for this it is important to understand the phenomena that get transfered from L1 to L2 (native language interference). Emotion is one of these. Linguistics research (Dewaele, 2010) has focused on the way emotions are encoded in different text types and in different languages. How to express emotion appropriately is related to the origin of the speaker (country, region), situational context in which social norms might be different (formal vs. informal setting), interlocutors (age, gender, social distance), topic.

As emotions are psychological constructions of cultural meaning, there may be a misfit between emotions and social context when individuals change cultural contexts or live two cultural models (Leersnyder et al., 2011). The use of emotions is considered both culture- and language-specific (Wierzbicka, 1994, 1999). We hypothesize that this leads to different emotion signals in writings in a second language, by authors with different native languages.

We test this hypothesis through multi-class classification of the L1 of the authors of essays written in L2 in different experimental set-ups that take into account proficiency levels and topics of the written essays. We encode emotion information using polarity and sentiment information from the NRC Word-Emotion Association Lexicon (NRC emotion lexicon) (Mohammad and Turney, 2013), taking into account not only the fine-grained (word-level) emotion information, but also general aspects of the written material (overall high- or low-emotion load). The results show that emotional information contributes to detecting the native language of the speaker.

2 Related Work

Caldwell-Harris (2014) shows that emotion usage depends on the language by focusing on differences in emotion usage in L1 and L2. The author states that there is a correlation between the usage of emotions and proficiency levels and the age a language is acquired.

While emotion-based features have been used in other NLP tasks, such as sentiment analysis (Sidorov et al., 2013), classification of documents into the corresponding emotion category (Wen and Wan, 2014), deception detection (Newman et al., 2003), among others, they are an underexplored area of second language writing.

Torney *et al.* (2012) use psycholinguistic features extracted by the Linguistic Inquiry and Word Count (LIWC) tool (Pennebaker et al., 2007) to identify the first language of an author, where emotion-based features are included as part of the feature vector, e.g., percentage of positive/negative emotion words. The LIWC feature set used in the paper also contains other types of features, e.g., personal concern categories (work, leisure), paralinguistic dimensions (assents, fillers,

nonfluencies), which obscure the contribution of the actual emotion features.

Rangel and Rosso (2013; 2016) investigate and confirm the hypothesis that the use of emotions depends on author's age and gender. The authors used a graph-based approach, where each node and edge were represented by the corresponding part-of-speech (POS) tag, then the representation was enriched with semantic information, emoticons, and with emotion information, which included polarity of words (polarity of common nouns, adjectives, adverbs or verbs in a sentiment lexicon) and emotionally charged words (replacing common nouns, adjectives, adverbs or verbs with the emotion information from the Spanish Emotion Lexicon (Sidorov et al., 2013)). The representation combining all the features described above was used with a SVM classifier.

Rangel and Rosso (2013; 2016) suggest that there are commonalities in the use of emotions across author age and gender. We examine the hypothesis that there are commonalities in the use of emotions in L2 across different L1s, suggested by the linguistic and psycholinguistic studies (Leersnyder et al., 2011; Wierzbicka, 1999). We test this by evaluating the impact of emotion-based features on classifying the L1 of the authors of essays written in L2.

3 Emotion features for NLI

The best performing features for NLI are word and character n-grams (Jarvis et al., 2013). They cover – and obscure – a wide range of phenomena, because language usage has multiple dimensions that can reveal information such as age, gender, cultural influences. In this study, we investigate the impact of words that have an emotion signal, since studies have shown that emotion is culture specific (Wierzbicka, 1994, 1999), and thus could be indicative of the native language of a speaker.

3.1 Datasets

We conduct experiments on two datasets commonly used in NLI research:

TOEFL11 (Blanchard et al., 2013): the ETS Corpus of Non-Native Written English (TOEFL11) contains 1,100 essays in English (avg. 348 tokens/essay) for each of the 11 L1s: Arabic (ARA), Chinese (CHI), French (FRE), German (GER), Hindi (HIN), Italian (ITA), Japanese (JPN), Korean (KOR), Spanish (SPA), Telugu (TEL), and

Turkish (TUR). The essays were written in response to eight different writing prompts, all of which appear in all 11 L1 groups. The dataset contains information regarding the proficiency level (low, medium, high) of the authors.

ICLE (Granger et al., 2009): the ICLEv2 dataset consists of essays written by highly-proficient non-native college-level students of English. We used a 7-language subset of the corpus normalized for topic and character encoding (Tetreault et al., 2012; Ionescu et al., 2014) to which we refer as ICLE. This subset contains 110 essays (avg. 747 tokens/essay after tokenization and removal of metadata) for each of the 7 languages: Bulgarian (BUL), Chinese (CHI), Czech (CZE), French (FRE), Japanese (JPN), Russian (RUS), and Spanish (SPA).

3.2 Experiment setup

We used the (pre-)tokenized version of TOEFL11 and tokenized ICLE with the Natural Language Toolkit (NLTK)[1] tokenizer. ICLE metadata was removed in pre-processing. Each essay was represented through the sets of features described below, using term frequency (tf) and the liblinear scikit-learn (Pedregosa et al., 2011) implementation of Support Vector Machines (SVM) with OvR (one vs. the rest) multi-class strategy. We report classification accuracy on 10-fold cross-validation experiments.

3.3 Features

3.3.1 Part-of-speech tags and function words

POS tag n-grams and function words (FWs) are considered core features in NLI research (Malmasi and Dras, 2015), not susceptible to topic bias, unlike word and character n-grams (Brooke and Hirst, 2011).

POS n-grams, n=1..3 POS features capture the morpho-syntactic patterns in a text, and are indicative of the L1, especially when used in combination with other types of features (Cimino and Dell'Orletta, 2017; Markov et al., 2017). POS tags were obtained with TreeTagger (Schmid, 1999), which uses the Penn Treebank tagset (36 tags).

Function words (FWs) n-grams, n=1..3 Function words clarify the relationships between the content-carrying elements of a sentence, and introduce syntactic structures like verbal complements,

---

[1]http://www.nltk.org

relative clauses, and questions (Smith and Witten, 1993). They are considered one of the most important stylometric features (Kestemont, 2014). The FW feature set consists of 318 English FWs from the scikit-learn package (Pedregosa et al., 2011). With respect to emotion features, FWs can appear as quantifiers, intensifiers (e.g., very good) or modify the emotion expressed in other ways.

3.3.2 Emotion words

We use the 14,182 emotion words and their associations with eight emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive) from the NRC emotion lexicon (Mohammad and Turney, 2013). Table 1 presents the emotion words statistics for our data.

TOEFL11				ICLE			
L1	No.	L1	%	L1	No.	L1	%
HIN	96,184	KOR	24.93	CZE	20,162	CHI	26.81
TEL	88,979	HIN	24.62	RUS	20,142	BUL	25.06
GER	88,268	CHI	24.32	BUL	18,939	JPN	24.74
CHI	87,486	TEL	24.19	SPA	17,187	RUS	24.72
TUR	83,945	JPN	24.15	CHI	16,794	FRE	23.88
KOR	82,878	TUR	23.90	FRE	16,750	CZE	23.81
FRE	82,454	FRE	23.30	JPN	16,234	SPA	23.33
SPA	81,497	GER	23.21				
ITA	75,339	ITA	23.16				
JPN	73,740	SPA	22.40				
ARA	69,156	ARA	21.91				

Table 1: Emotion words statistics (absolute number and frequency) sorted from the highest to the lowest.

Before committing to analyzing emotion features, we want to test whether emotion-loaded words have any impact on the NLI task. The bag-of-words (BoW) representation covers a variety of phenomena, without distinguishing them and giving us insight into their individual impact on the task. We represent our data using BoW variations – including and excluding words that have an emotional dimension. To verify that the effect in classification is not just due to a smaller feature set, we match the BoW size by removing a selection of random words. Table 2 presents the 10-fold cross-validation results (accuracy, %) on the TOEFL11 and ICLE datasets, when using emotion words and random words of such that the BoW representations have the same size, as well as the results when excluding emotion words and the random words.[2]

The results in Table 2 show that emotion words have higher impact on classification accuracy than random words when evaluated in isolation. Moreover, the accuracy drop is higher when excluding

[2]Random words accuracy was calculated as average over five experiments with five different sets of random words.

Features	TOEFL11		ICLE	
	Acc., %	No.	Acc., %	No.
BoW	68.65	61,339	80.65	20,032
Random words	36.15	8,187	70.21	6,465
Emotion words	46.75	8,187	72.86	6,465
BoW w/o random words	66.68	53,152	76.83	13,567
BoW w/o emotion words	63.11	53,152	75.19	13,567

Table 2: Performance of emotion words.

emotion words from the BoW approach than when excluding random words, confirming that emotion is a useful dimension for L1 classification, and not just an effect of having additional features.

3.3.3 Emotion features

Having confirmed that due to cultural identity and linguistic habits of an author's native language, we can distinguish the L1 of the author of an essay, we proceed with a deeper analysis, for which we build two types of emotion features.

Emotion polarity features (*emoP*) In the NRC emotion lexicon, binary associations are provided for each emotion word for 8 emotions (anger, fear, anticipation, trust, surprise, sadness, joy, or disgust) and two sentiments (negative or positive) – e.g., *good* = "0100101011". This representation is used as a categorial feature (not a 10-dimensional binary vector). It performed best compared to other ways of encoding the emotion information we tried, e.g., using a 10-dimensional binary vector or excluding the sentiment information.

The *emoP* features are added to the POS and to POS & FW representations: the phrase *This is very good* is represented through POS & emoP unigrams as 'DT', 'VBZ', 'RB', 'JJ-0100101011', or as 3-grams 'DT_VBZ_RB', VBZ_RB_JJ-0100101011', and as POS & FW & emoP 3-grams as 'This_is_very', 'is_very_JJ-0100101011'.

Emotion load features (*emoL*) Speakers of different L1s may use a higher or lower number of emotionally charged words than speakers of other L1s, reflecting cultural customs or linguistic habits of the respective cultures. We modeled this information using three types of emotion load features: (i) two binary features, *emoL* (binary) that capture whether an essay has a high or low emotional load: (a) we compute the average ratio of emotion words in all essays in each dataset: for TOEFL11 this was 0.236 and for ICLE 0.246; (b) if the ratio of emotion words in an essay was higher/lower than the average, assigned it a "highly-emotional"/"low-emotional" feature. We used this representation

to examine whether the polarity as such is informative. We also used more fine-grained *emoL* features: (ii) the ratio of the emotion words in each essay as a numeric feature (1 feature, *emoL* (1)), and (iii) the ratio of each emotion/sentiment in each essay (10 numeric features: 8 emotions and 2 sentiments, *emoL* (10)). Overall, three different types of *emoL* features are examined.

4 Results and Discussion

Following previous studies on NLI (Markov et al., 2018) and author profiling (Rangel and Rosso, 2016), we provide the results when adding emotion-based features to POS tag feature set. We also experiment with POS and FW feature sets similarly to, e.g., (Malmasi and Dras, 2015).

The 10-fold cross-validation results in terms of accuracy (%) on the TOEFL11 and ICLE datasets for POS and POS & FW n-gram (n = 1–3) representations are shown in Tables 3 and 4, respectively. The number of features (No.) is included. Statistically significant gains/drops according to McNemar's statistical significance test (McNemar, 1947) with $\alpha < 0.05$ are marked with '*'.

The experimental results show that emotion features, in particular the *emoP* features, significantly contribute to the results for all the considered settings, indicating that different cultures (as defined by the authors' L1) have different emotion word usage. It is very interesting to note that despite being very general, the three types of *emoL* features – 13 features that characterize the emotional load of a document – also improve the results in the majority of settings, including when combined with the *emoP* features. This supports the hypothesis that some cultures use a bigger or smaller emotional vocabulary. More fine grained emotional load features could improve the results further.

To explore whether emotion usage depends on specific topics, we conducted experiments for the topics in the TOEFL11 dataset (Table 5).[3] The improvement brought by the emotion-based features does seem to depend on the topic, as some topics more naturally elicit emotional reactions. The highest improvements were achieved for P5 (car usage) and P7 (young vs. old people comparison). When combined with the POS & FW representation, emotion-based features are less helpful (not

---

[3] We did not conduct this experiment on the ICLE dataset, since it has a higher number of topics, with a fewer number of documents per topic, which would not allow us to learn informative topic-specific models.

statistically significant improvements) for the topics discussing traveling (P1), ideas vs. facts (P3), and education (P4). Overall, adding emotion-based features to POS and POS & FW representations leads to accuracy improvement for all the topics present in the dataset.

The ability to choose the proper words to express oneself increases with the proficiency level. From this perspective, identifying the L1 of authors of essays in L2 using emotion words information should be performed with better results. On the other hand, we expect other linguistic characteristics to become closer to a native L2 speaker, and thus make identifying L1 harder. We experiment with L1 classification separating the data based on the three different proficiency levels in TOEFL11. The results are included in Table 6. With respect to the emotion features, medium and high proficiency levels have a much better performance. As postulated above, this could be explained by the different ability of the L1 speakers to choose the words that express closely the message and nuances they wish to convey.

5 Conclusions

We investigated the hypothesis that the use of emotions is indicative of an author's native language. We used two types of emotion-based features – one that captures the types of sentiments expressed, the other captures the frequency of emotion words in documents. We expected these features to capture cultural characteristics and linguistic habits from the authors' L1. The fact that adding these features to POS and function word n-grams leads to improvements in predicting a text's author's native language leads us to conclude that emotion characteristics from a native language are "imported" into the production of L2.

The overall goal of this paper was to understand the influence of various facets of L1 speakers' language and culture on their acquisition (and production) of L2. These influences from L1 are not under the author's conscious control, and it is very interesting to understand their nature. Emotion is one of these. The fact that we explore the use of emotions on learner corpora ("controlled environment"), with a specific task and specific requirement – and a (implied, not specifically requested) more neutral style – should probably lower the effect of emotional influences from the L1 and its culture. From that point of view, it is even more remarkable that such an effect is detected.

Features	TOEFL11		ICLE	
	Acc., %	No.	Acc., %	No.
POS 1–3-grams (baseline)	40.16	17,483	62.86	11,755
POS 1–3-grams + emoL (binary)	40.60	17,485	62.86	11,757
POS 1–3-grams + emoL (1)	40.19	17,484	62.86	11,756
POS 1–3-grams + emoL (10)	40.41	17,493	62.99	17,765
POS 1–3-grams + emoL (binary) + emoL (1) + emoL (10)	40.65	17,496	62.60	11,768
Difference:	**0.49***		**−0.26**	
POS 1–3-grams + emoP	50.36	216,090	67.66	90,920
Difference:	**10.20***		**4.80***	
POS 1–3-grams + emotion-based features	50.28	216,103	67.79	90,933
Difference (with POS 1–3 + emoP):	**−0.08**		**0.13**	
Difference (with baseline):	**10.12***		**4.93***	

Table 3: 10-fold cross-validation accuracy for POS 1–3-grams combined with emotion-based features. '*' marks statistically significant differences.

Features	TOEFL11		ICLE	
	Acc., %	No.	Acc., %	No.
POS 1–3-grams	40.16	17,483	62.86	11,755
POS & FW 1–3-grams (baseline)	64.06	411,599	74.42	138,170
Difference:	**23.90***		**11.56***	
POS & FW 1–3-grams + emoL (binary)	64.10	411,601	74.42	138,172
POS & FW 1–3-grams + emoL (1)	64.10	411,600	74.42	138,171
POS & FW 1–3-grams + emoL (10)	64.09	411,609	74.42	138,180
POS & FW 1–3-grams + emoL (binary) + emoL (1) + emoL (10)	64.13	411,612	74.42	138,183
Difference:	**0.07**		**0.00**	
POS & FW 1–3-grams + emoP	67.73	880,595	77.92	268,605
Difference:	**3.67***		**3.50***	
POS & FW 1–3-grams + emotion-based features	67.85	880,608	78.31	268,618
Difference (with POS & FW 1–3 + emoP):	**0.12**		**0.39**	
Difference (with baseline):	**3.79***		**3.89***	

Table 4: 10-fold cross-validation accuracy for POS & FW 1–3-grams combined with emotion-based features. '*' marks statistically significant differences.

	P0	P1	P2	P3	P4	P5	P6	P7
POS 1–3-grams	33.74	39.26	38.54	39.89	42.40	38.29	42.08	38.15
POS 1–3-grams + emotion-based features	41.14	47.34	44.79	46.02	49.12	49.55	49.01	47.19
Difference:	**7.40***	**8.08***	**6.25***	**6.13***	**6.72***	**11.26***	**6.93***	**9.04***
POS & FW 1–3-grams	50.54	56.54	53.28	55.62	60.34	56.84	57.79	55.31
POS & FW 1–3-grams + emotion-based features	53.18	57.66	56.56	57.04	62.28	62.40	61.46	58.66
Difference:	**2.64***	**1.12**	**3.28***	**1.42**	**1.94**	**5.56***	**3.67***	**3.35***
No. of emotion words:	99,606	75,308	116,795	118,427	122,741	129,837	107,924	139,288
Ratio:	0.213	0.239	0.222	0.226	0.238	0.239	0.243	0.274

Table 5: 10-fold cross-validation accuracy for each topic in the TOEFL11 dataset. '*' marks statistically significant differences.

	Low		Medium		High	
	Acc., %	No.	Acc., %	No.	Acc., %	No.
POS 1–3-grams	41.10	9,751	43.07	15,334	34.65	14,454
POS 1–3-grams + emotion-based features	44.56	51,108	52.64	152,059	42.58	136,783
Difference:	**3.46***		**9.57***	L	**7.93***	
POS & FW 1–3-grams	52.40	91,340	66.52	288,658	54.25	242,880
POS & FW 1–3-grams + emotion-based features	54.13	155,725	69.09	585,083	57.20	491,342
Difference:	**1.73**		**2.57***		**2.95***	
No. of emotion words:	62,223		475,665		372,025	
Ratio:	0.228		0.235		0.242	

Table 6: 10-fold cross-validation accuracy for each proficiency level. '*' marks statistically significant differences.

References

Daniel Blanchard, Joel Tetreault, Derrick Higgins, Aoife Cahill, and Martin Chodorow. 2013. TOEFL11: A corpus of non-native English. *ETS Research Report Series*, 2013(2):i–15.

Julian Brooke and Graeme Hirst. 2011. Native language detection with 'cheap' learner corpora. In *Proceedings of the Conference of Learner Corpus Research*, pages 37–47, Louvain-la-Neuve, Belgium. Presses universitaires de Louvain.

Catherine Caldwell-Harris. 2014. Emotionality differences between a native and foreign language: Theoretical implications. *Frontiers in Psychology*, 5(1055).

Andrea Cimino and Felice Dell'Orletta. 2017. Stacked sentence-document classifier approach for improving native language identification. In *Proceedings of the 12th Workshop on Building Educational Applications Using NLP*, pages 430–437, Copenhagen, Denmark. ACL.

Jean-Marc Dewaele. 2010. *Emotions in Multiple Languages*. Basingstoke: Palgrave Macmillan.

Sylviane Granger, Estelle Dagneaux, Fanny Meunier, and Magali Paquot. 2009. *International Corpus of Learner English v2 (ICLE)*. Presses Universitaires de Louvain, Louvain-la-Neuve, Belgium.

Radu Tudor Ionescu, Marius Popescu, and Aoife Cahill. 2014. Can characters reveal your native language? A language-independent approach to native language identification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 1363–1373, Doha, Qatar. ACL.

Scott Jarvis, Yves Bestgen, and Steve Pepper. 2013. Maximizing classification accuracy in native language identification. In *Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 111–118, Atlanta, GA, USA. ACL.

Mike Kestemont. 2014. Function words in authorship attribution. From black magic to theory? In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature*, pages 59–66, Gothenburg, Sweden. ACL.

Jozefien De Leersnyder, Batja Mesquita, and Heejung S. Kim. 2011. Where do my emotions belong? a study of immigrants' emotional acculturation. *Personality and Social Psychology Bulletin*, 37(4):451–463.

Shervin Malmasi and Mark Dras. 2015. Multilingual native language identification. *Natural Language Engineering*, 23(2):163–215.

Ilia Markov, Lingzhen Chen, Carlo Strapparava, and Grigori Sidorov. 2017. CIC-FBK approach to native language identification. In *Proceedings of the 12th Workshop on Building Educational Applications Using NLP*, pages 374–381, Copenhagen, Denmark. ACL.

Ilia Markov, Vivi Nastase, and Carlo Strapparava. 2018. Punctuation as native language interference. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3456–3466, Santa Fe, New Mexico, USA. The COLING 2018 Organizing Committee.

Quinn McNemar. 1947. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157.

Saif Mohammad and Peter Turney. 2013. Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29:436–465.

Matthew Newman, James Pennebaker, Diane Berry, and Jane Richards. 2003. Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin*, 29(5).

Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

James Pennebaker, Roger Booth, and Martha Francis. 2007. *Linguistic Inquiry and Word Count: LIWC2007*. Austin, TX: LIWC.net.

Francisco Rangel and Paolo Rosso. 2013. On the identification of emotions and authors' gender in Facebook comments on the basis of their writing style. In *Proceedings of the First International Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and perspectives from AI*, volume 1096, pages 34–46, Torino, Italy. CEUR-WS.org.

Francisco Rangel and Paolo Rosso. 2016. On the impact of emotions on author profiling. *Information Processing & Management*, 52(1):74–92.

Helmut Schmid. 1999. *Improvements In Part-of-Speech Tagging With an Application to German*. Springer.

Grigori Sidorov, Sabino Miranda-Jiménez, Francisco Viveros-Jiménez, Alexander Gelbukh, Noé Castro-Sánchez, Francisco Velásquez, Ismael Díaz-Rangel, Sergio Suárez-Guerra, Alejandro Treviño, and Juan Gordon. 2013. Empirical study of machine learning based approach for opinion mining in tweets. In *Proceedings of the Mexican International Conference on Artificial Intelligence*, volume 7629, pages 1–14, San Luis Potosí. Mexico. Springer.

Tony C. Smith and Ian H. Witten. 1993. Language inference from function words. Working papers, https://hdl.handle.net/10289/9927.

Joel Tetreault, Daniel Blanchard, Aoife Cahill, and Martin Chodorow. 2012. Native tongues, lost and found: Resources and empirical evaluations in native language identification. In *Proceedings of the 24th International Conference on Computational Linguistics*, pages 2585–2602, Mumbai, India. The COLING 2012 Organizing Committee.

Rosemary Torney, Peter Vamplew, and John Yearwood. 2012. Using psycholinguistic features for profiling first language of authors. *Journal of the Association for Information Science and Technology*, 63(6).

Shiyang Wen and Xiaojun Wan. 2014. Emotion classification in microblog texts using class sequential rules. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*, pages 187–193, Quebec, Canada. AAAI Press.

Anna Wierzbicka. 1994. Emotion, language, and cultural scripts. *Emotion and culture: Empirical studies of mutual influence*, pages 133–196.

Anna Wierzbicka. 1999. *Emotions across languages and cultures: Diversity and universals*. Cambridge University Press.