Does 'well-being' translate on Twitter?

Laura K. Smith[1], Salvatore Giorgi[1], Rishi Solanki[1], Johannes C. Eichstaedt[1], H. Andrew Schwartz[2], Muhammad Abdul-Mageed[3], Anneke Buffone[1] and Lyle H. Ungar[1]

[1]University of Pennsylvania, [2]Stony Brook University, [3]University of British Columbia

Objectives

- Does psychological well-being translate across English and Spanish on Twitter?
- Is meaningful cultural information lost in translation?

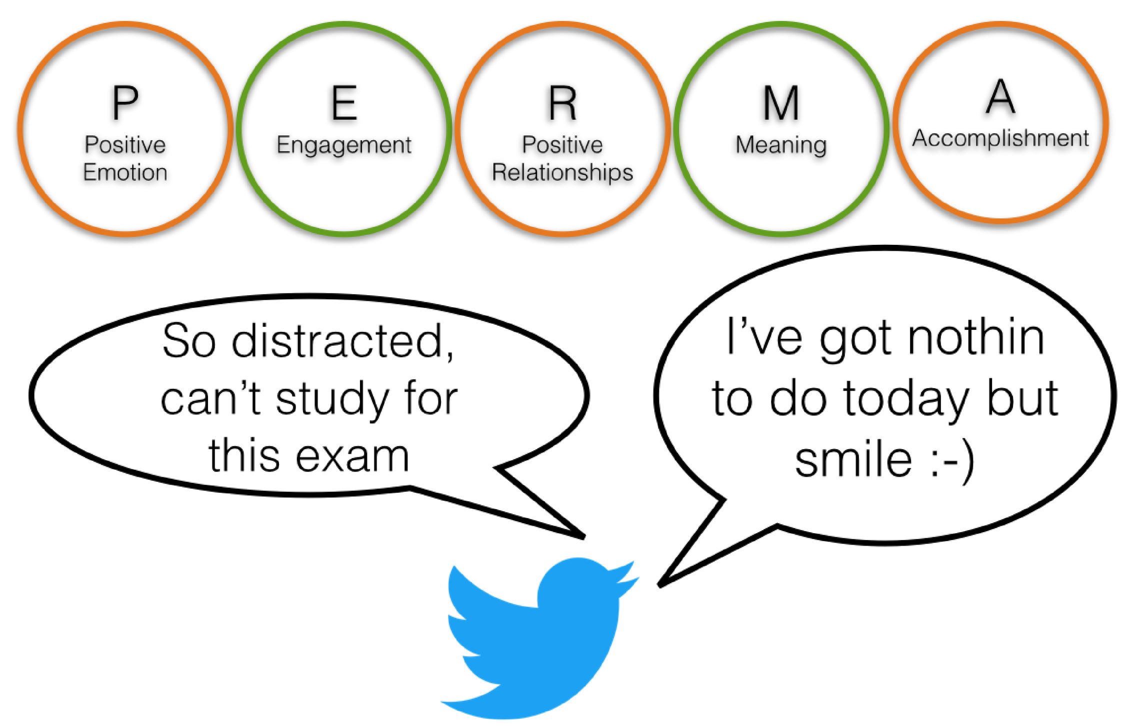
Specifically:

- Can *resources* built in English or Spanish be translated and applied to text in the other language?
- Can *text* in English or Spanish be automatically translated to the other language in order to apply resources developed in that language?

Introduction

- Improvements in SMT systems allow sentiment in resource-poor languages to be assessed by translating text into a resource-rich language such as English and applying an English sentiment model [1]
- Approach is economical and efficient
- Not clear how much culturally specific information and accuracy are lost in translation
- Less research has examined the translation of sentiment on social media
- Research has not focused on translating subjective well-being

Well-Being

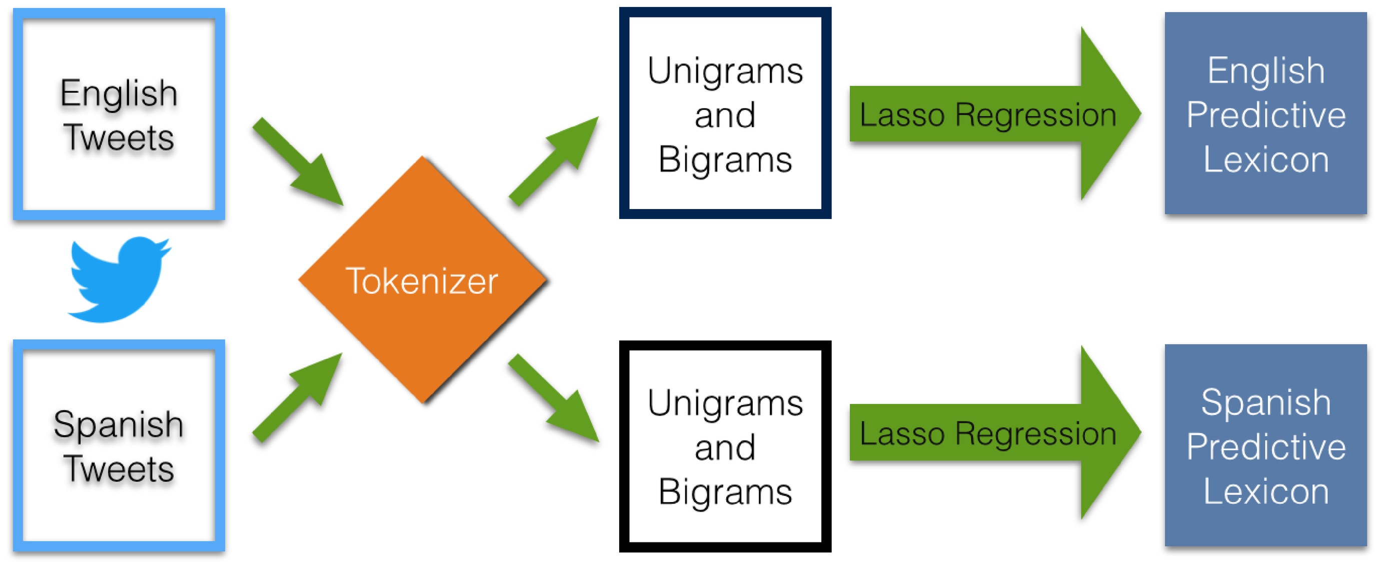


- PERMA is a five-dimensional model of well-being [2]
- **P** dimension maps relatively cleanly onto traditional conceptions of sentiment
- PERMA includes social and cognitive components which may be expressed with more variation across languages and cultures
- PERMA model already developed using Facebook data [3]

Data and Annotations

- Amazon's Mechanical Turk (MTurk) used to annotate 5,000 random English Tweets
- CrowdFlower used to annotate 5,000 random Spanish Tweets
- Separate annotation tasks set up for each of the 10 PERMA components
- Workers asked to indicate "to what extent does this message express" the construct on a scale from 1 to 7

Methods



- Tweets were tokenzied using an emoticon-aware tokenizer
- Vocabularies of ~5,000 unigrams/bigrams in English and Spanish
- Regression model used to predict the average annotation score
- Models were then transformed into a predictive lexicon [4]

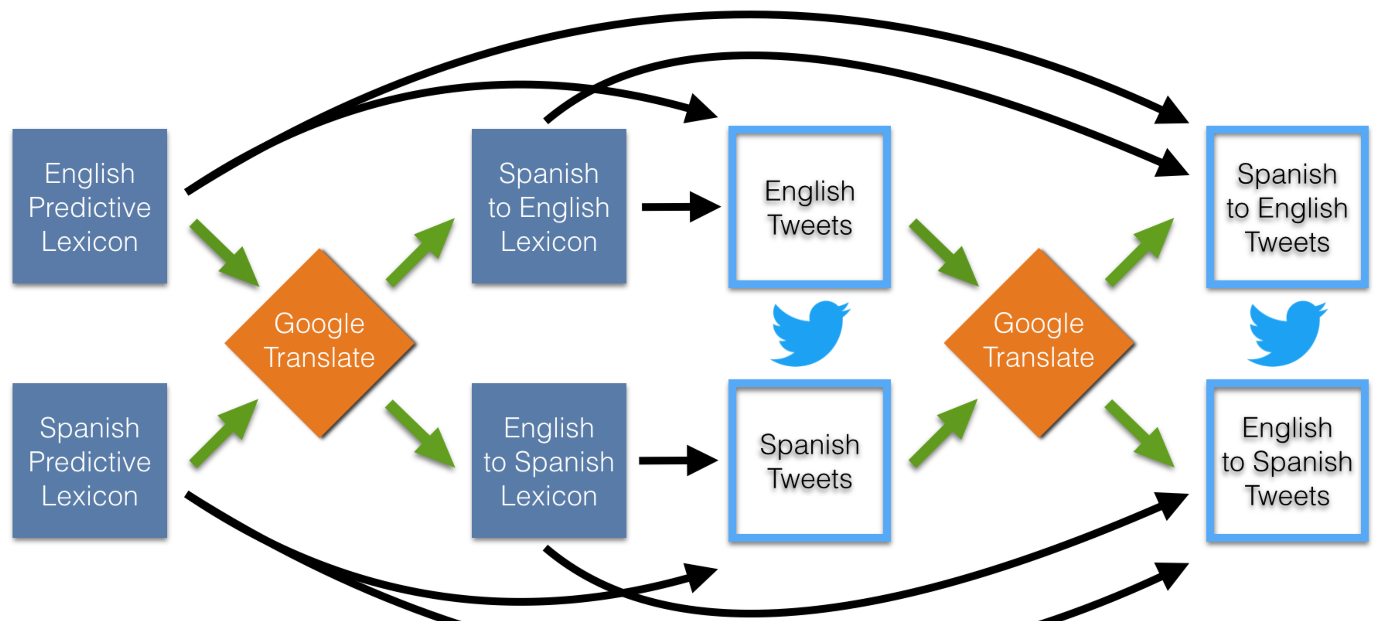
Model	r
Spanish	0.36
English	0.36

Table 1: Performance as measured by Pearson r correlation averaged over the 10 positive/negative PERMA components using 10-fold cross validation

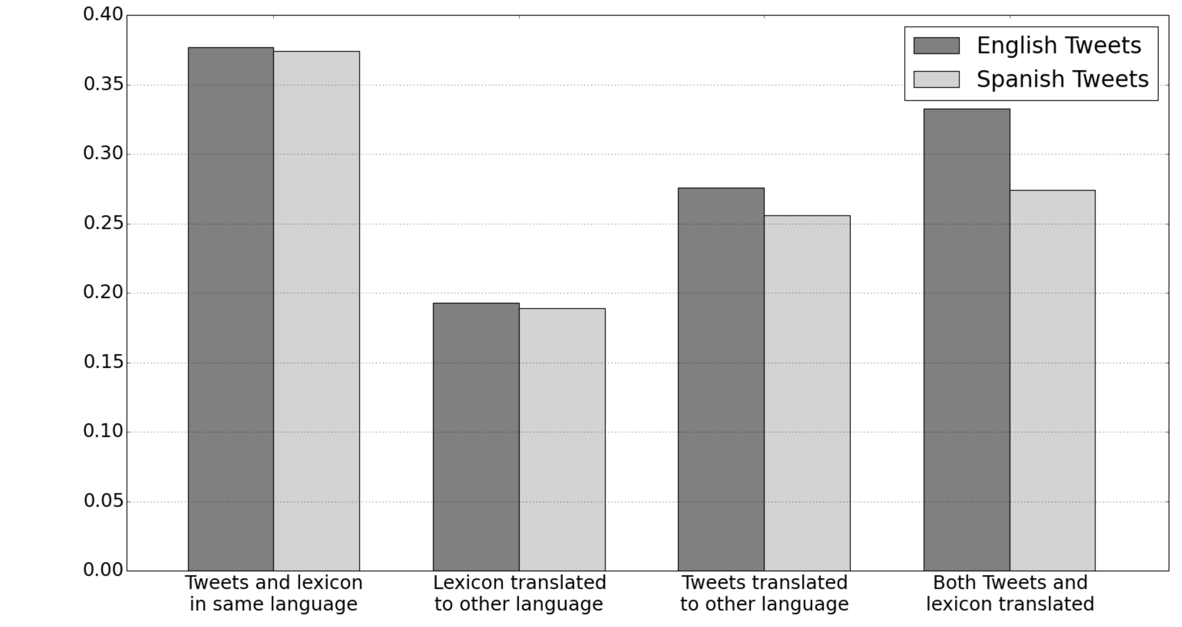
PERMA	Top Weighted Words
+P	excited, beautiful, awesome, blessed, so happy
-P	kill, pissed, hate, sad, annoying
+E	night, life, im, i'm, playing
-E	boring, bored, feel like, sleep, alone
+R	awesome, proud, love u, happy, love you
-R	tired of, f*ck, f*cking, one to, a**hole
+M	god, jesus, string, life, . proud
-M	bored, myself, tired, f*ck, sleep
+A	made, finished, almost, pretty, . http://
-A	wish i, myself, dnt, everything, gonna

Table 2: Top five weighted words for each PERMA category in the English predictive lexicon

Results



- Lexica built on 80% of the messages and then evaluated on the remaining 20%
- Performance reported as Pearson r correlation between ground-truth annotations and predicted lexica scores
- Scores averaged over the 10 PERMA components



Error Analysis

source lang	correct trans	missing terms	opp sign	weight diff
English	83%	81%	0.5%	6.9%
Spanish	74%	91%	0.0%	4.8%

Table 3: Summary of translation errors

- *correct trans*: percentage of correctly translated words
- *missing terms*: percentage of correct translations not appearing in the other model
- *opp sign*: percentage of terms whose signs switched between models
- *weight diff*: percentage of terms with significant weight differences between models

PERMA	term	weight (en)	weight (es)	% chg
POS_M (en)	mundo* (world)	0.42	-0.18	143
NEG_A (en)	odio** (hate)	0.29	2.19	87
NEG_M (en)	nadie*** (no one)	0.23	0.24	4.2
NEG_R (es)	sad** (triste)	1.70	0.0012	100
NEG_P (es)	hate*** (odio)	1.81	1.75	3.3

Table 4: Examples of specific errors

- Error types denoted by asterisks: * denotes a change in sign, ** denotes the largest change in weight and *** denotes the smallest change in weight per source model
- Language listed under each PERMA category is the language of the source model that was translated
- The *% chg* column is percentage change relative to the larger weight

Conclusion

- Source language models applied to the source language Tweets performed best
- Translating a single piece (model or Tweets) resulted in a decrease in performance
- Translating both (model and Tweets) performed better than translating one piece
- Manually correcting translation errors did not improve model performance
- Suggests that meaningful cultural information was lost

References

[1] Saif M Mohammad, Mohammad Salameh, and Svetlana Kiritchenko. How translation alters sentiment. *Journal of Artificial Intelligence Research*, 55:95–130, 2016.

[2] Martin EP Seligman. *Flourish.* Free Press, New York, NY, 2011.

[3] H Andrew Schwartz, Maarten Sap, Margaret L Kern, Johannes C Eichstaedt, Adam Kapelner, Megha Agrawal, et al. Predicting individual well-being through the language of social media. In *Biocomputing 2016: Proceedings of the Pacific Symposium*, pages 516–527, 2016.

[4] Maarten Sap, Greg Park, Johannes C Eichstaedt, Margaret L Kern, David J Stillwell, Michal Kosinski, et al. Developing age and gender predictive lexica over social media. In *Proceedings of the 2014 Conference on Empirical Methods In Natural Language Processing*, EMNLP, pages 1146–1151, 2014.

Acknowledgements

The authors acknowledge support from the Templeton Religion Trust (grant TRT-0048) and Bioibérica.

Contact Information

- Web: http://www.wwbp.org
- Email: lasm@sas.upenn.edu, sgiorgi@sas.upenn.edu