Scoring Mechanism to Quantify Different Emotions

Madhumitha Sivaraj
Deep Data Lab
Rutgers University
Piscataway, NJ 08854
madhu.sivaraj@rutgers.edu

Abstract

Empathy is a central part of our lives and leads to communication and cooperation, transcending social interactions, behaviors, and well-being. An AI system with a capacity for empathy could provide more natural interactions while making judgments that take moods and feelings into account. I built a scoring mechanism to quantify different emotions based on data-driven methods, including word vectors and corpora. Inspired by Linguistic Inquiry and Word Count, I created a system to score various Amazon reviews on a scale of 0 to 1 for all relevant categories from the comprehensive list in the default LIWC2015 dictionary.

1 Introduction

How words are used by individuals in their every day lives may convey or provide information regarding their values, beliefs, modes of thought, social interactions, and personalities. Despite advances in computer science research and artificial intelligence, machines lack the common sense to understand human language. The latest natural language processing (NLP) techniques have taken great strides to mitigate such issues that hinder the research and exploration in these ambiguous areas of study. However, the hard problem of artificial intelligence is that machines are unable to identify the various emotional, cognitive, and structural components present in verbal and written speech, so there is huge room for improvement for current methods. My goal is to create an alternative to existing mechanisms and develop a scoring mechanism based on data containing various emotional, cognitive, and structural components.

Language is a medium for exchanging information and communicating with others around us. Machines need to understand the intricacies of language and how people interact and make use of both verbal and written speech. In today's society, intelligence voice assistants such as Amazon Alexa, Apple Siri, and Google Assistant are expanding rapidly with new products entering the market and a strong emphasis on both email and voice user interfaces. These agents use natural language processing to interpret user text or voice input and execute commands and respond via synthesized voices. Users can ask their assistant questions, control home automation devices, and media playback via voice, and manage other basic tasks such as email, to-do lists, and calendars with verbal commands. However, in their current state, most intelligent voice assistants powered by artificial intelligence have room for improvement in regards to integrating empathy and emotional content. The way engineers design interactions with AI systems and the results they provide should be thoughtfully considered, and in the future, the responsibility for designing artificial empathy[1]. An AI system with the capacity for empathy could provide more natural interactions while making judgments that take our mood or feelings into account.

The way they deliver information could be modulated to appropriately respond to an individual's current emotional state, resulting in more natural experience, but this is no easy task. Currently, machine learning algorithms are quite good at recognizing some physiological or behavioral cues that are associated with specific emotional states. However, there are a variety of sophisticated and

complex theories of emotion and how they're produced. It is not as simple as recognizing someone's tone of voice, posture, or facial expression. In-depth knowledge of empathy is required to build upon AI systems' understanding of different emotional nuances and improving algorithms to better understand the emotional cues of individual users. This would allow virtual assistant systems to go beyond only measuring the surface level of expressions.

Linguistic Inquiry and Word Count (LIWC)[2] is one of the leading textual analysis software tools which was designed to provide an efficient and effective method for studying the various emotional, cognitive, and structural components present in individuals' verbal and written speech samples. The tool uses machine learning to read a text, categorize it by word type, and assign a LIWC-derived score to the word based on its category. In 2020, Facebook built and open-sourced BlenderBot[3], the largest-ever open-domain chatbot which outperforms other chatbots in terms of engagement and also feels more human, according to human evaluators. The culmination of years of research in conversational AI, this is the first chatbot to blend a diverse set of conversational skills — including empathy, knowledge, and personality — together in one system. Amazon uses their Neural text-to-speech (NTTS) model[4] to enable more natural-sounding speech by wrapping Alexa's response with the appropriate SSML tag ('excited' or 'disappointed') and the level of intensity with which the emotion should be applied to the response ('low', 'medium' or 'high'). However, these models are not yet designed to integrate the emotional nuances and the aforementioned complex theories of emotion.

My approach is to use the comprehensive list of the default LIWC2015 dictionary categories, scales, and sample scale words and GloVe word vectors to find nearest neighbors of the LIWC example words. I map words to LIWC dictionary categories and derive a score to arbitrary paragraphs on a scale of 0 to 1 for all relevant LIWC categories. One limitation that can be improved upon is that the scoring mechanism does not factor context from surrounding words in the input.

2 Related Work

Linguistic Inquiry and Word Count (LIWC2015) is a simple software tool developed by researchers at UT Austin led by James W. Pennebaker that merely counts words through a lightweight processor, but its dictionaries and its use in over two decades of empirical research have made it one of the leading software programs. To get a sense of its capabilities, it helps to review the contents of its main built-in dictionary, which includes summary dimensions of the ingested text, punctuation marks, function words, affect terms (including sentiment analysis), social terms, cognitive process terms, perceptual processes, biological processes, human drives, time orientation, relativity, personal concerns, and informal language. On the surface, the software offers simple counts in these various categories, but these dictionaries are based on psychometric and other types of research. The dictionaries have been used by researchers in various fields (more on this later).

The LIWC2015 Dictionary is the heart of the text analysis strategy. The default LIWC2015 Dictionary is composed of almost 6,400 words, word stems, and select emoticons. Each dictionary entry additionally defines one or more word categories or subdictionaries. For example, the word cried is part of five-word categories: sadness, negative emotion, overall affect, verbs, and past focus. Hence, if the word cried is found in the target text, each of these five subdictionary scale scores will be incremented. As in this example, many of the LIWC2015 categories are arranged hierarchically. All sadness words, by definition, belong to the broader "negative emotion" category, as well as the "overall affect words" category. Each of the default LIWC2015 categories is composed of a list of dictionary words that define that scale.

Figure 1 provides a comprehensive list of the default LIWC2015 dictionary categories, scales, sample scale words, and relevant scale word counts. I use certain data in this table provided in the LIWC2015 Manual including the categories and the example words for each category. This information is incorporated into my scoring mechanism when mapping each word in the input to the LIWC dictionary categories.

I was also inspired by Empath[5], a tool that can generate and validate new lexical categories on demand from a small set of seed terms (like "bleed" and "punch" to generate the category violence). Empath draws connotations between words and phrases by deep learning a neural embedding across more than 1.8 billion words of modern fiction. Given a small set of seed words that characterize a category, Empath uses its neural embedding to discover new related terms, then validates the

Category	Abbrev	Examples	Words in category	Internal Consistency (Uncorrected α)	Internal Consistency (Corrected α)
Friends	friend	buddy, neighbor	95	.20	.60
Female references	female	girl, her, mom	124	.53	.87
Male references	male	boy, his, dad	116	.52	.87
Cognitive processes	cogproc	cause, know, ought	797	.65	.92
Insight	insight	think, know	259	.47	.84
Causation	cause	because, effect	135	.26	.67
Discrepancy	discrep	should, would	83	.34	.76
Tentative	tentat	maybe, perhaps	178	.44	.83
Certainty	certain	always, never	113	.31	.73
Differentiation	differ	hasn't, but, else	81	.38	.78
Perceptual processes	percept	look, heard, feeling	436	.17	.55
See	see	view, saw, seen	126	.46	.84
Hear	hear	listen, hearing	93	.27	.69
Feel	feel	feels, touch	128	.24	.65
Biological processes	bio	eat, blood, pain	748	.29	.71
Body	body	cheek, hands, spit	215	.52	.87
Health	health	clinic, flu, pill	294	.09	.37
Sexual	sexual	horny, love, incest	131	.37	.78
Ingestion	ingest	dish, eat, pizza	184	.67	.92
Drives	drives		1103	.39	.80
Affiliation	affiliation	ally, friend, social	248	.40	.80
Achievement	achieve	win, success, better	213	.41	.81
Power	power	superior, bully	518	.35	.76
Reward	reward	take, prize, benefit	120	.27	.69
Risk	risk	danger, doubt	103	.26	.68
Time orientations	TimeOrient				
Past focus	focuspast	ago, did, talked	341	.23	.64
Present focus	focuspresent	today, is, now	424	.24	.66
Future focus	focusfuture	may, will, soon	97	.26	.68
Relativity	relativ	area, bend, exit	974	.50	.86
Motion	motion	arrive, car, go	325	.36	.77
Space	space	down, in, thin	360	.45	.83
Time	time	end, until, season	310	.39	.79
Personal concerns					
Work	work	job, majors, xerox	444	.69	.93
Leisure	leisure	cook, chat, movie	296	.50	.86
Home	home	kitchen, landlord	100	.46	.83
Money	money	audit, cash, owe	226	.60	.90
Religion	relig	altar, church	174	.64	.91
Death	death	bury, coffin, kill	74	.39	.79
Informal language	informal		380	.46	.84
Swear words	swear	fuck, damn, shit	131	.45	.83
Netspeak	netspeak	btw, lol, thx	209	.42	.82
Assent	assent	agree, OK, yes	36	.10	.39
Nonfluencies	nonflu	er, hm, umm	19	.27	.69
Fillers	filler	Imean, youknow	14	.06	.27

Figure 1: LIWC2015 Output Variable Information.

category with a crowd-powered filter. Empath also analyzes text across 200 built-in, pre-validated categories generated from common topics in their web dataset, like neglect, government, and social media. Empath's data-driven, human validated categories were highly incorporated with similar categories in LIWC. Empath can generate new lexical categories and analyze text over 200 built-in human-validated categories.

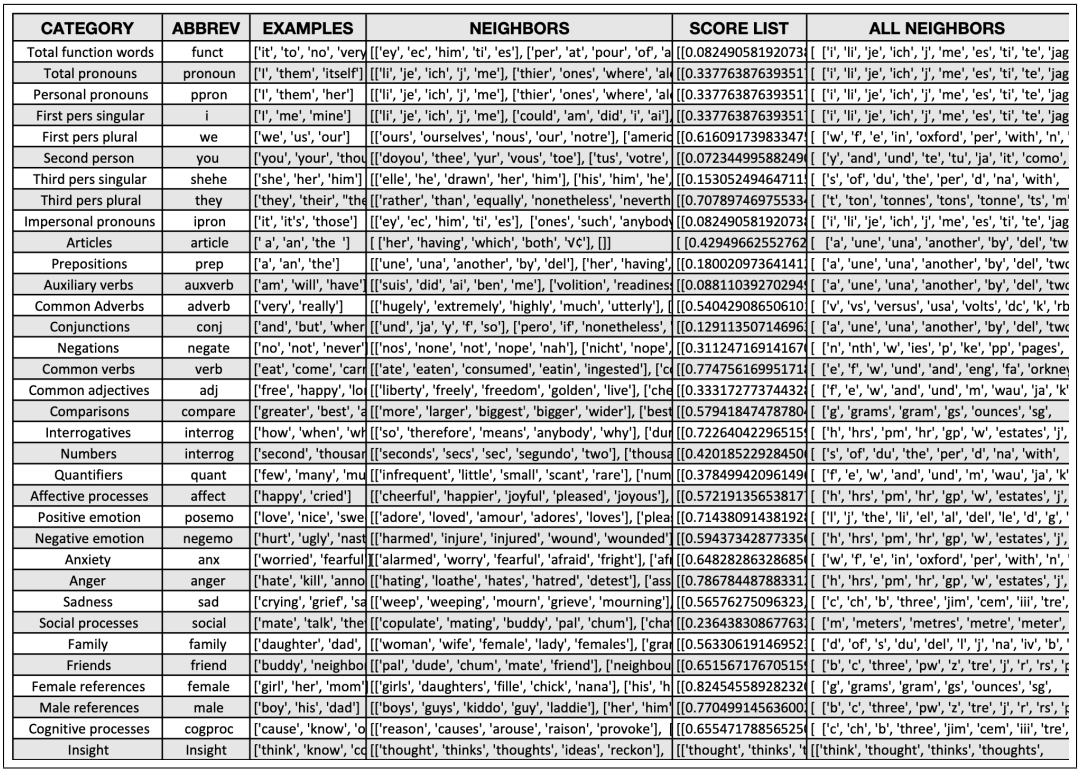
3 Overview

My goal is to build a LIWC-inspired scoring mechanism to quantify different emotions based on data-driven methods, including word vectors and corpora. I will be able to score arbitrary texts similar to how LIWC does on a scale of 0 to 1 for all relevant categories from the comprehensive list of the default LIWC2015 dictionary.

3.1 Implementation

I created a simple CSV file where each line was a LIWC category followed by the list of examples words for that category in the LIWC2015 manual. Then for each LIWC category, I found the nearest neighbors to each LIWC example word using GLoVE word embeddings. I then created a data frame where I store the categories and their abbreviation, their LIWC example words, the nearest neighbors

for each example words, and the sum of similarity scores. The figure below illustrates a portion of the data frame containing the data for the first 34 LIWC categories.



CATEGORY	ABBREV	EXAMPLES	NEIGHBORS	SCORE LIST	ALL NEIGHBORS
Total function words	funct	['it', 'to', 'no', 'very	[['ey', 'ec', 'him', 'ti', 'es'], ['per', 'at', 'pour', 'of', 'a	[[0.08249058192073	[['i', 'li', 'je', 'ich', 'j', 'me', 'es', 'ti', 'te', 'jag
Total pronouns	pronoun	['I', 'them', 'itself']	[['li', 'je', 'ich', 'j', 'me'], ['thier', 'ones', 'where', 'al	[[0.33776387639351	[['i', 'li', 'je', 'ich', 'j', 'me', 'es', 'ti', 'te', 'jag
Personal pronouns	ppron	['I', 'them', 'her']	[['li', 'je', 'ich', 'j', 'me'], ['thier', 'ones', 'where', 'al	[[0.33776387639351	[['i', 'li', 'je', 'ich', 'j', 'me', 'es', 'ti', 'te', 'jag
First pers singular	i	['I', 'me', 'mine']	[['li', 'je', 'ich', 'j', 'me'], ['could', 'am', 'did', 'i', 'ai'],	[[0.33776387639351	[['i', 'li', 'je', 'ich', 'j', 'me', 'es', 'ti', 'te', 'jag
First pers plural	we	['we', 'us', 'our']	[['ours', 'ourselves', 'nous', 'our', 'notre'], ['americ	[[0.61609173983347	[['w', 'f', 'e', 'in', 'oxford', 'per', 'with', 'n', '
Second person	you	['you', 'your', 'thou	[['doyou', 'thee', 'yur', 'vous', 'toe'], ['tus', 'votre',	[[0.07234499588249	[['y', 'and', 'und', 'te', 'tu', 'ja', 'it', 'como',
Third pers singular	shehe	['she', 'her', 'him']	[['elle', 'he', 'drawn', 'her', 'him'], ['his', 'him', 'he',	[[0.15305249464711	[['s', 'of', 'du', 'the', 'per', 'd', 'na', 'with',
Third pers plural	they	['they', 'their', 'the	[['rather', 'than', 'equally', 'nonetheless', 'neverth	[[0.70789746975533	[['t', 'ton', 'tonnes', 'tons', 'tonne', 'ts', 'm'
Impersonal pronouns	ipron	['it', 'it's', 'those']	[['ey', 'ec', 'him', 'ti', 'es'], ['ones', 'such', 'anybod	[[0.08249058192073	[['i', 'li', 'je', 'ich', 'j', 'me', 'es', 'ti', 'te', 'jag
Articles	article	[' a', 'an', 'the ']	[['her', 'having', 'which', 'both', 'V¢'], []]	[[0.42949662552762	[['a', 'une', 'una', 'another', 'by', 'del', 'tw
Prepositions	prep	['a', 'an', 'the']	[['une', 'una', 'another', 'by', 'del'], ['her', 'having',	[[0.18002097364141	[['a', 'une', 'una', 'another', 'by', 'del', 'two
Auxiliary verbs	auxverb	['am', 'will', 'have']	[['suis', 'did', 'ai', 'ben', 'me'], ['volition', 'readines	[[0.08811039270294	[['a', 'une', 'una', 'another', 'by', 'del', 'two
Common Adverbs	adverb	['very', 'really']	[['hugely', 'extremely', 'highly', 'much', 'utterly'], [	[[0.54042908650610	[['v', 'vs', 'versus', 'usa', 'volts', 'dc', 'k', 'rb
Conjunctions	conj	['and', 'but', 'wher	[['und', 'ja', 'y', 'f', 'so'], ['pero', 'if', 'nonetheless',	[[0.12911350714696	[['a', 'une', 'una', 'another', 'by', 'del', 'two
Negations	negate	['no', 'not', 'never']	[['nos', 'none', 'not', 'nope', 'nah'], ['nicht', 'nope',	[[0.31124716914167	[['n', 'nth', 'w', 'ies', 'p', 'ke', 'pp', 'pages',
Common verbs	verb	['eat', 'come', 'carr	[['ate', 'eaten', 'consumed', 'eatin', 'ingested'], ['c	[[0.77475616995171	[['e', 'f', 'w', 'und', 'and', 'eng', 'fa', 'orkney
Common adjectives	adj	['free', 'happy', 'lo	[['liberty', 'freely', 'freedom', 'golden', 'live'], ['che	[[0.33317277374432	[['f', 'e', 'w', 'and', 'und', 'm', 'wau', 'ja', 'k'
Comparisons	compare	['greater', 'best', 'a	[['more', 'larger', 'biggest', 'bigger', 'wider'], ['best	[[0.57941847478780	[['g', 'grams', 'gram', 'gs', 'ounces', 'sg',
Interrogatives	interrog	['how', 'when', 'wh	[['so', 'therefore', 'means', 'anybody', 'why'], ['dur	[[0.72264042965155	[['h', 'hrs', 'pm', 'hr', 'gp', 'w', 'estates', 'j',
Numbers	interrog	['second', 'thousar	[['seconds', 'secs', 'sec', 'segundo', 'two'], ['thousa	[[0.42018522928450	[['s', 'of', 'du', 'the', 'per', 'd', 'na', 'with',
Quantifiers	quant	['few', 'many', 'mu	[['infrequent', 'little', 'small', 'scant', 'rare'], ['num	[[0.37849942096149	[['f', 'e', 'w', 'and', 'und', 'm', 'wau', 'ja', 'k'
Affective processes	affect	['happy', 'cried']	[['cheerful', 'happier', 'joyful', 'pleased', 'joyous'],	[[0.57219135653817	[['h', 'hrs', 'pm', 'hr', 'gp', 'w', 'estates', 'j',
Positive emotion	posemo	['love', 'nice', 'swe	[['adore', 'loved', 'amour', 'adores', 'loves'], ['plea	[[0.71438091438192	[['l', 'j', 'the', 'li', 'el', 'al', 'del', 'le', 'd', 'g', '
Negative emotion	negemo	['hurt', 'ugly', 'nast	[['harmed', 'injure', 'injured', 'wound', 'wounded']	[[0.59437342877335	[['h', 'hrs', 'pm', 'hr', 'gp', 'w', 'estates', 'j',
Anxiety	anx	['worried', 'fearful	[['alarmed', 'worry', 'fearful', 'afraid', 'fright'], ['af	[[0.64828286328685	[['w', 'f', 'e', 'in', 'oxford', 'per', 'with', 'n', '
Anger	anger	['hate', 'kill', 'anno	[['hating', 'loathe', 'hates', 'hatred', 'detest'], ['ass	[[0.78678448788331	[['h', 'hrs', 'pm', 'hr', 'gp', 'w', 'estates', 'j',
Sadness	sad	['crying', 'grief', 'sa	[['weep', 'weeping', 'mourn', 'grieve', 'mourning'],	[[0.56576275096323	[['c', 'ch', 'b', 'three', 'jim', 'cem', 'iii', 'tre',
Social processes	social	['mate', 'talk', 'the	[['copulate', 'mating', 'buddy', 'pal', 'chum'], ['cha	[[0.23643830867763	[['m', 'meters', 'metres', 'metre', 'meter',
Family	family	['daughter', 'dad',	[['woman', 'wife', 'female', 'lady', 'females'], ['gra	[[0.56330619146952	[['d', 'of', 's', 'du', 'del', 'll', 'j', 'na', 'iv', 'b', '
Friends	friend	['buddy', 'neighbo	[['pal', 'dude', 'chum', 'mate', 'friend'], ['neighbou	[[0.65156717670515	[['b', 'c', 'three', 'pw', 'z', 'tre', 'j', 'r', 'rs', 'g
Female references	female	['girl', 'her', 'mom'	[['girls', 'daughters', 'fille', 'chick', 'nana'], ['his', 'h	[[0.82454558928232	[['g', 'grams', 'gram', 'gs', 'ounces', 'sg',
Male references	male	['boy', 'his', 'dad']	[['boys', 'guys', 'kiddo', 'guy', 'laddie'], ['her', 'him'	[[0.77049914563600	[['b', 'c', 'three', 'pw', 'z', 'tre', 'j', 'r', 'rs', 'g
Cognitive processes	cogproc	['cause', 'know', 'o	[['reason', 'causes', 'arouse', 'raison', 'provoke'], [	[[0.65547178856525	[['c', 'ch', 'b', 'three', 'jim', 'cem', 'iii', 'tre',
Insight	Insight	['think', 'know', 'co	[['thought', 'thinks', 'thoughts', 'ideas', 'reckon'],	[['thought', 'thinks', 't	[['think', 'thought', 'thinks', 'thoughts',

Figure 2: Data frame including neighbors and scores.

Then, I read in an input text and determine which LIWC category each word in the input corresponds to by traversing the initial table and checking if each word is a LIWC example word or one of the nearest neighbors for each word. I use not only the word but also four stemming and lemmatization algorithms that provide multiple variations of the word so I can better traverse the initial data frame and map words to their corresponding categories. Some words can map to multiple categories so I prioritize categories in the order of example words, the 10 nearest neighbors, and then all neighbors.

After determining which categories each word from the input belongs to, I score each word among the example words and consider the highest score. I average the scores from the input words for all words that belong to a specific category and assign a derived score to the input based on its category. I repeat this process for all categories represented in the input text.

4 Results

To evaluate my I scored 20 Amazon reviews (10 5-star and 10 1-star) on a set of 10 products using the scoring mechanism to analyze trends in their reviews and scores. See Figure S1 (5-Star reviews) and Figure S2 (1-Star reviews) in Appendix for tables containing information about these review texts and their derived scores.

The results revealed that more words in all 20 reviews, both the 5-star and 1-star, mapped to the category, "Total function words". This is understandable, and quite frankly, expected as the example words for this category were "it, to, no, very". The top categories for both 5-star and 1-star reviews included similar categories including conjunctions, prepositions, total pronouns, impersonal pronouns, and articles. Words that fit under this category serve grammatical and structural purposes. These function words often do not carry meaning.

According to James Pennebaker, function words account for less than one-tenth of 1 percent of the vocabulary but make up almost 60 percent of the words used. Hence, it is important to look at categories that serve meanings that are represented less than functional word categories.

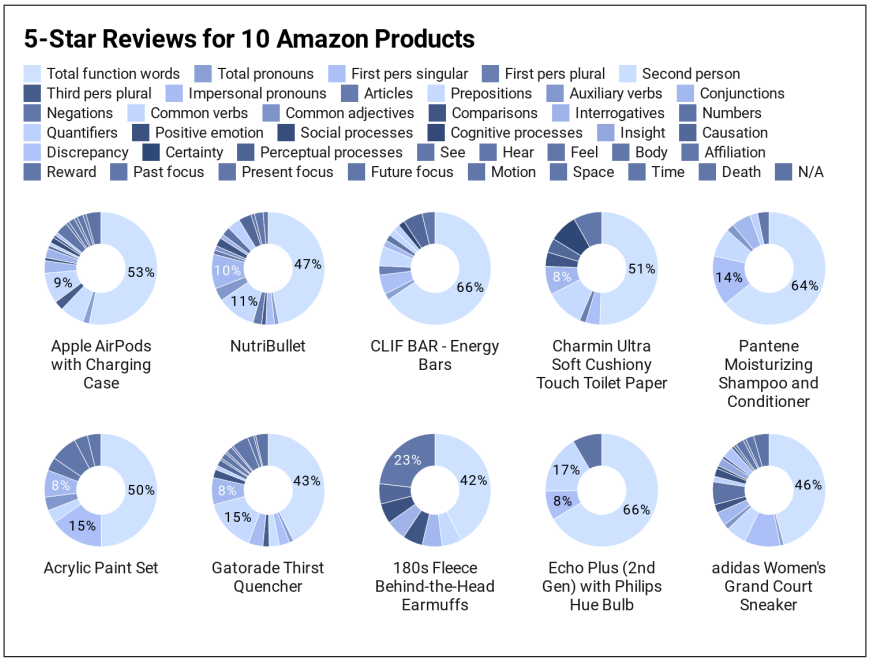


Figure 3: Illustrates category representations in 5-star Amazon reviews.

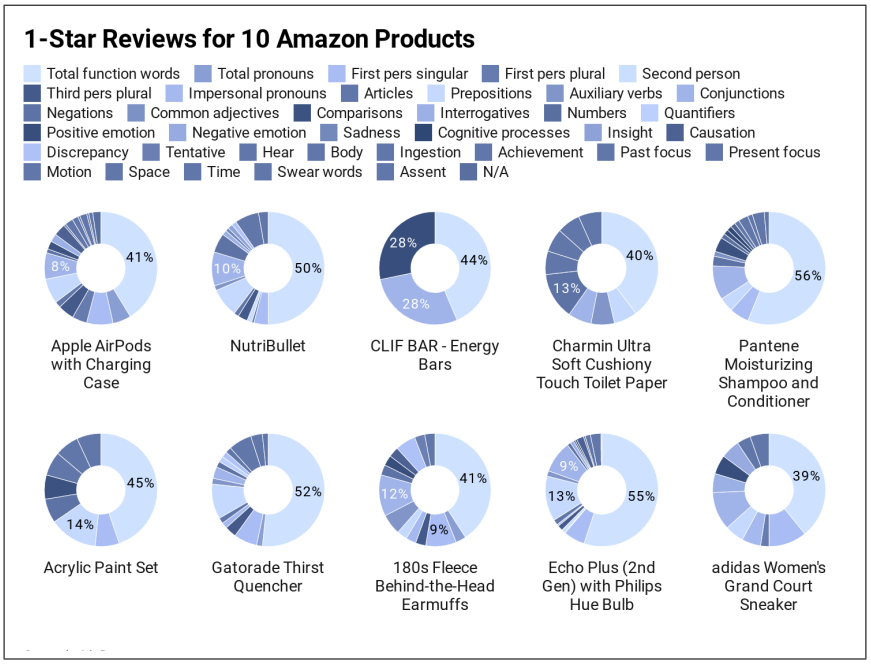


Figure 4: Illustrates category representations in 1-star Amazon reviews.

Figures 3 and 4 show that categories with content words appear far less than functional words. The content word categories that appear most in 5-star Amazon reviews include hear, space, social processes, numbers, and future focus. This is justifiable as positive reviews may include information about how the customer uses the product on a daily basis, how its price compares to other models, and

whether or not the customer intends to continue purchasing it. Meanwhile, causation, comparisons, motion, negation, and discrepancy appear more in 1-star reviews. These results are expected as negative reviews tend to complain about a product and bring up issues and discrepancies in products and orders, often comparing the product received to the way the product was advertised on Amazon.com.

5 Conclusion

I produced a scoring mechanism inspired by LIWC when researching to quantify different emotions based on data-driven methods, including word vectors and corpora. As tech giants work to expand the capabilities of AI agents powered by NLP, it is important to improve natural language understanding and natural language generation techniques such that AI systems have a capacity for empathy. This will allow products like intelligent voice assistants to be more natural in their interactions and factor the user's mood and feelings when generating responses and executing commands. The mechanism mapped words to LIWC dictionary categories to score paragraphs on a scale of 0 to 1 for multiple categories. I scored 20 Amazon reviews (10 5-star and 10 1-star) on the same set of products to analyze trends in their reviews and scores and found that categories with content words appear far less than functional words.

6 Future Work

One limitation of the scoring mechanism was that it did not consider surrounding words for context when assigning derived scores in the input text. In the future, I will integrate linguistic features such as word usage and context through an n-gram language model and psycholinguistic attributes via LIWC to see if this would allow for more accurate results. I will also explore using machine learning to predict how the relevant words are.

7 Acknowledgments

I thank Dr. Gerard DeMelo, Director of the Deep Data Lab, Professor in the Department of Computer Science, and Affiliate Member of the Center for Cognitive Science at Rutgers University for his guidance and support.

References

[1] Esat Dedezade. 2019. Jobs of the future: teaching empathy to artificial intelligence.

[2] James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psychometric properties of LIWC2015. Austin, TX: University of Texas at Austin.

[3] Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, and Jason Weston: "Recipes for building an open-domain chatbot", 2020; In *arXiv preprint arXiv:2004.1363*.

[4] Catherine Gao. 2019. Use New Alexa Emotions and Speaking Styles to Create a More Natural and Intuitive Voice Experience.

[5] 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*.

Appendix

Figure S1 displays 5-Star Reviews and Figure S2 displays 1-Star Reviews for 10 products along with their respective results from the scoring mechanism. For each category represented in the Amazon review text, a corresponding score is attached to it.

PRODUCT	5 STAR REVIEW (TEXT)	5 STAR REVIEW (SCORE)
Apple AirPods with Charging Case	These AirPods are amazing they automatically play audio as soon as you put them in your ears and pause when you take them out. A simple double-tap during music listening will skip forward. To adjust the volume, change the song, make a call, or even get directions, just say Hey Siri to activate your favorite personal assistant. Plus, when youre on a call or talking to Siri, an additional accelerometer works with dual beamforming microphones to filter out background noise and ensure that your voice is transmitted with clarity and consistency. Additionally, they deliver five hours of listening time on a single charge, and theyre made to keep up with you thanks to a charging case that holds multiple additional charges for more than 24 hours of listening time. Just 15 minutes in the case gives you three hours of listening to time or up to two hours of talk time. I would highly recommend it to anyone looking to buy	[('Impersonal pronouns', 0.006211180124223602), ('Total function words', 0.3710216546676471), ('Present focus', 0.012422360248447204), ('Third pers plural', 0.018633540372670808), ('Prepositions', 0.06109350212044 2), ('Future focus', 0.006211180124223602), ('Second person', 0.048576977060678386), ('Total pronouns', 0.012422360248447204), ('Space', 0.012422360248447204), ('Conjunctions', 0.024844720496894408), ('Interrogatives', 0.018633540372670808), ('Reward', 0.006211180124223602), ('Hear', 0.024844720496894408), ('Auxiliary verbs', 0.006211180124223602), ('Time', 0.006211180124223602), ('Social processes', 0.011201012732048913), ('Articles', 0.006211180124223602), ('Quantifiers', 0.006211180124223602), ('Causation', 0.006211180124223602), ('Comparisons', 0.006211180124223602), ('Numbers', 0.006211180124223602), ('First pers singular', 0.006211180124223602), ('Discrepancy', 0.006211180124223602), ('Perceptual processes', 0.004632739438096 84)]
NutriBullet	This is my third NB purchase and by far the most impressive product. The first two were 600 and 900 watts, which lasted about 3 years each when the motors burned out. I assume this was because of filling too close to the lines. This model has several improvements, including multiple speeds and functions, an actual blender, plus the 32 and 24 oz cups for smoothies; it is quieter and more high tech. The old style cup had rubber gaskets that were inset and fouled if not very carefully cleaned by removing them. The new cups have eliminated this problem. I expect to enjoy this NB more and for it to last longer than my first two.	[('Total function words', 0.32572424341974016), ('Present focus', 0.01709401709401096), ('Conjunctions', 0.06837606837606838), ('Prepositions', 0.07692307692307693), ('Quantifiers', 0.02564102564102564), ('Numbers', 0.01709401709401096), ('Articles', 0.01709401709401096), ('Interrogatives', 0.008547008547008548), ('Motion', 0.010243483491488658), ('First pers singular', 0.01709401709401096), ('Causation', 0.02564102564102564), ('Auxiliary verbs', 0.02564102564102564), ('Affiliation', 0.006891297399653198), ('Comparisons', 0.01709401709401096), ('Negations', 0.008547008547008548), ('Total pronouns', 0.008547008547008548), ('Common adjectives', 0.008547008547008548), ('Third pers plural', 0.008547008547008548)]
CLIF BAR - Energy Bars	shipped fast, product was in perfect order and shape. flavor GREAT price GOOD Popular with work out crew lots of flavors the guys loved flavors the guys loved Low sugar high protein	[('Total function words', 0.36854884863779364), ('Space', 0.03125), ('Conjunctions', 0.03125), ('Positive emotion', 0.09375), ('Prepositions', 0.0625), ('Male references', 0.0625)]
Charmin Ultra Soft Cushiony Touch Toilet Paper	I've been using this toilet paper for several years now. I order a case each month at the much-cheaper Amazon Subscribe & Save price. This paper is thick, very absorbent, doesnt contain any scents or dyed that cause any irritation for me or my family. One case of 18 huge rolls lasts our family with 2 bathrooms for 1 whole month. I've tried most all other brands and this one lasts the longest of any of them. I pray that Amazon never stops carrying this product.	[('Total function words', 0.40842370000446765), ('First pers plural', 0.016221099412348516), ('Causation', 0.03488372093023256), ('Quantifiers', 0.011627906976744186), ('Present focus', 0.023255813953488372), ('First pers singular', 0.03488372093023256), ('Prepositions', 0.03488372093023256), ('Cognitive processes', 0.011627906976744186), ('Conjunctions', 0.011627906976744186), ('Total pronouns', 0.011627906976744186), ('Negations', 0.011627906976744186), ('Common verbs', 0.011627906976744186)]
Pantene Moisturizing Shampoo and Conditioner	My fav, go-to shampoo and conditioner for years. I occasionally add another brand but always go back to Pantene. Products 4x the price don‚Äôt work as well! I‚Äôve used Purology, Tea Tree Paul Mitchell, Aveda, etc, always go back to this!	[('Total function words', 0.3026626173355583), ('Conjunctions', 0.04878048780487805), ('Causation', 0.024390243902439025), ('First pers singular', 0.024390243902439025), ('Prepositions', 0.06583881258987 68), ('Certainty', 0.04878048780487805), ('Motion', 0.04878048780487805), ('Comparisons', 0.024390243902439025), ('First pers plural', 0.010823015738245125)]
Acrylic Paint Set	24 small tubes (12ml, or .4 ounces) of a wide variety of colors. I am just trying to get back into some creative projects, and I thought I would refresh my stash of paint. If I knew I would be getting into bigger projects, I would buy larger containers. When I saw that this was a set of small tubes, I thought it would be perfect for me as I didn't want to any to go to waste. It is enough to start dabbling again without spending a lot of money to buy larger amounts. Comes with 3 basic paint brushes, too. All of the tubes in my package look good. I pressed all of them, and they feel soft. I opened up half dozen, and those were all fine. Hopefully remainder are good. I think this is a nice kit to have if you want to dabble, or as an add on to a themed gift.	[('Total function words', 0.36942761724950995), ('Quantifiers', 0.012738853503184714), ('Prepositions', 0.04712600402783374), ('First pers singular', 0.08280254777070063), ('Auxiliary verbs', 0.012738853503184714), ('Space', 0.019108280254777707), ('Conjunctions', 0.03184713375796178), ('33', 0.019108280254777707), ('Discrepancy', 0.025477070063694 28), ('Comparisons', 0.019108280254777707), ('Interrogatives', 0.006369426751592357), ('See', 0.006369426751592357), ('Causation', 0.006369426751592357), ('Motion', 0.006369426751592357), ('Present focus', 0.019108280254777707), ('Social processes', 0.006369426751592357), ('Common verbs', 0.004112816280312351), ('Perceptual processes', 0.006369426751592357), ('Positive emotion', 0.019108280254777707), ('Total pronouns', 0.008329725576610015), ('Third pers plural', 0.006369426751592357), ('Impersonal pronouns', 0.006369426751592357), ('Second person', 0.006369426751592357), ('Articles', 0.006369426751592357)]
Gatorade Thirst Quencher	I depend on this drink to keep me hydrated and my body systems in balance. Since I am diabetic, this is a low-calorie, tasty way to keep me going. And this is my most favorite Gatorade flavor	[('First pers singular', 0.10810810810810811), ('Total function words', 0.3495042195046213 5), ('Conjunctions', 0.05405405405405406), ('Space', 0.027027027027027 03), ('Past focus', 0.027027027027027 03), ('Auxiliary verbs', 0.027027027027027 03), ('Present focus', 0.05405405405405406), ('Prepositions', 0.027027027027027 03), ('Motion', 0.027027027027027 03)]
180s Fleece Behind-the-Head Earmuffs	I bought a pair of these in D,C, in November from a street vendor. For a south Texan, D.C. in November was unbearable; it was cold with incessant wind and little sunshine. After I purchased these from the Pentagon City metro stop vendor, my last week in the Capital City was more than bearable. I spent hours on the Mall going from Smithsonian to Smithsonian. I spent hours in Arlington Cemetery as well as walking around Georgetown and the canal district. They are easy to don, although it takes both hands to put on. Doffing these ear muffs is even easier: use one hand to pull back the 'yoke' at the back of your head, fold them using the inherent twist and collapse feature with one hand, and they easily fit in a winter coat pocket without too much bulk. Combine these behind the head muffs with a tight weave winter cap and touch screen winter gloves, and you'll stay as warm as possible without feeling like Randy from 'A Christmas Story' These have been such a great addition to my Winter wardrobe, even in South Texas. that I stuffed them in stockings for my family, albeit after the holiday season. You won't be disappointed. And, as an added bonus for all your gotta-keep-my-coiffure-perfect friends and family members, these muffs sit on the ears, but the suspension rides on the occipital bone. If the temperatures are pleasant enough, you won't muss your hair. But the temperatures require a stocking cap, ymmv, Highly recommended.	[('First pers singular', 0.019762845849802372), ('Total function words', 0.2834458840149 15), ('Prepositions', 0.10077848450958396), ('Impersonal pronouns', 0.027667984818972332), ('Space', 0.03162055335968379), ('Causation', 0.011857707509881422), ('Conjunctions', 0.051383399209486 17), ('Quantifiers', 0.007905138339920948), ('Comparisons', 0.015810276679841896), ('Third pers plural', 0.011857707509881422), ('Motion', 0.003952569169960474), ('Death', 0.003952569169960474), ('Present focus', 0.011857707509881422), ('Reward', 0.002763779550035 22), ('Third pers singular', 0.003952569169960474), ('Body', 0.011857707509881422), ('Time', 0.011857707509881422), ('First pers plural', 0.003406394983050 1644), ('Second person', 0.019762845849802372), ('Total pronouns', 0.007905138339920948), ('Feel', 0.003952569169960474), ('Auxiliary verbs', 0.003952569169960474), ('Articles', 0.003952569169960474), ('Assent', 0.001791789477885 6226), ('Affiliation', 0.003952569169960474), ('Positive emotion', 0.003952569169960474)]
Echo Plus (2nd Gen) with Philips Hue Bulb	Short & Sweet - For what it is, this device is amazing. Performance is quick, audio is great, and the built-in Zigbee hub works quite well.	[('Total function words', 0.28017566563357466), ('Positive emotion', 0.038461538461538464), ('Causation', 0.038461538461538464), ('Interrogatives', 0.038461538461538464), ('Present focus', 0.15384615384615385), ('Conjunctions', 0.038461538461538464), ('Prepositions', 0.038461538461538464), ('Comparisons', 0.038461538461538464)]
adidas Women's Grand Court Sneaker	Great running shoe right out of the box! Highly recommended it.....love this shoe. I run 40-50 miles a week.	[('Total function words', 0.41506595939157915), ('Prepositions', 0.10526315789473684), ('First pers singular', 0.05263157894736842), ('Numbers', 0.05263157894736842)]

Figure S1: Derived scores for 5-Star Reviews.

PRODUCT	1 STAR REVIEW (TEXT)	1 STAR REVIEW (SCORE)
Apple AirPods with Charging Case	We bought a brand new set of AirPods for 159. After using them for a week, I was listening to them and the right air pod went dead. I figured, maybe I just need to charge them. So when I got home I charged them, and the next day I went to use them again and the right air bud was still dead. So I did some research, tried to reset them, and then I couldnt reconnect them to my phone at all. Once I contacted Apple, they tried to give us the run around. They wanted us to let them fully die, which makes sense, so we did that. Then, when they still didnt work we wanted to exchange the faulty pair for a new set (BECAUSE WE PAID 159 DOLLARS AND ONLY USED THEM FOR A WEEK). Keep in mind, they were taken care of, not dropped, no water damage, they were expensive so they were treated delicately. Apple wanted a 180 deposit to get a new set, which is MORE THAN WHAT I ORIGINALLY PAID FOR. I decided to contact Amazon, and they suggested contacting Apple, but once I explained what was going, on Amazon offered to make things right and send out a replacement. Never had an issue with Amazon customer service, but Apple was extremely disappointing. I wont purchase these again.	[('First pers plural', 0.0323254936990254O4), ('Total function words', 0.31475411575015133), ('Prepositions', 0.054465100510835814), ('Causation', 0.026548672566371 68), ('Comparisons', 0.017699115044247787), ('Total pronouns', 0.03982300884955752), ('First pers singular', 0.057522123893805 31), ('Hear', 0.004424778761061947), ('Conjunctions', 0.057522123893805 31), ('Motion', 0.015396594352023887), ('Time', 0.008849557522123894), ('Tentative', 0.004424778761061947), ('Interrogatives', 0.017699115044247787), ('Past focus', 0.013274336283185 84), ('Differentiation', 0.004424778761061947), ('Ingestion', 0.017699115044247787), ('Third pers plural', 0.035398230088495575), ('Articles', 0.013274336283185 84), ('Space', 0.004424778761061947), ('Negations', 0.008849557522123894), ('Present focus', 0.004424778761061947), ('Auxiliary verbs', 0.004424778761061947), ('Impersonal pronouns', 0.004424778761061947)]
NutriBullet	Very loud! I didn't think noise would ever be an issue but this is WAY louder than any blender I ‚Äôve owned in the past. The base is also very heavy and suctions to surfaces, so not ideal if you are constantly getting out/putting away (I personally do not like appliances sitting out on the counter when not in use). The smaller sized cups are so tall with just a tiny blade at the bottom, I can never get it to blend my ingredients - so basically they are useless. Sadly will be throwing this away and getting something else ASAP.	[('Total function words', 0.361022922058207), ('First pers singular', 0.03), ('33', 0.01), ('Discrepancy', 0.01), ('Articles', 0.01), ('Conjunctions', 0.07), ('Present focus', 0.05), ('Third pers plural', 0.02), ('Space', 0.02), ('Prepositions', 0.06), ('Negations', 0.04), ('Second person', 0.01), ('Interrogatives', 0.01), ('First pers plural', 0.004668033812497192), ('Sadness', 0.00713904794391362 25), ('Auxiliary verbs', 0.01)]
CLIF BAR - Energy Bars	Tastes terrible and super sweet	[('Total function words', 0.3084179131240047), ('Conjunctions', 0.2), ('Positive emotion', 0.2)]
Charmin Ultra Soft Cushiony Touch Toilet Paper	Rolls are way too narrow and not worth the cost. Basucally Family Mega...translates into better have child size hands. Never buy again!!	[('Total function words', 0.26911576188186204), ('Present focus', 0.045454545454545456), ('Conjunctions', 0.045454545454545456), ('Negations', 0.09090909090909091), ('Prepositions', 0.045454545454545456), ('Space', 0.045454545454545456), ('Achievement', 0.045454545454545456), ('Auxiliary verbs', 0.045454545454545456), ('Body', 0.045454545454545456)]
Pantene Moisturizing Shampoo and Conditioner	Never been a fan of pantene but read some good reviews and decided to try this shampoo and conditioner set especially because it said it was sulfate free... At first it made my hair shiny that was after the first wash.. Then After the second wash my hair started to fall out in chunks and It makes me mad cause I was already growing it out I was left with my hair being all thinned out and i lost about 2 inches in length..please do not buy this shampoo and conditioner its crap! There are better options out there and cheaper!	[('Negations', 0.019801980198019802), ('Total function words', 0.3987497667644932), ('Prepositions', 0.02970297029702 97), ('Conjunctions', 0.06930693069306931), ('Positive emotion', 0.009900990099009901), ('Causation', 0.009900990099009901), ('Common adjectives', 0.009900990099009901), ('Comparisons', 0.02970297029702 97), ('Numbers', 0.009900990099009901), ('Space', 0.02441857336725205), ('First pers singular', 0.039603960396039604), ('Cognitive processes', 0.009900990099009901), ('Present focus', 0.019801980198019802), ('Motion', 0.009900990099009901), ('Swear words', 0.009900990099009901), ('Achievement', 0.009900990099009901)]
Acrylic Paint Set	Half of the colors are dry. I did not realized it after the return policy expired. Very disappointed.	[('Total function words', 0.35958968527664015), ('Prepositions', 0.1111111111111111), ('Present focus', 0.055555555555555555), ('First pers singular', 0.055555555555555555), ('Past focus', 0.055555555555555555), ('Negations', 0.055555555555555555), ('Comparisons', 0.055555555555555555), ('Time', 0.055555555555555555)]
Gatorade Thirst Quencher	There is nothing classic about this vile wretched drink. I purchased these hoping to recapture the classic refreshing, revitalizing, experience from my athletic Highschool days. I am beside myself on the lack of transparency in the advertisement (no mention of small bottles in the title) and i find it misleading to call this product classic. It is clearly smaller than the standard classic gatorade and the flavor is something I would not wish on my worst enemy. This is nothing less than an atrocity.	[('Total function words', 0.3689023967496324 3), ('Present focus', 0.047619047619047616), ('First pers singular', 0.047619047619047616), ('Impersonal pronouns', 0.011904761904761904), ('Prepositions', 0.07142857142857142), ('Time', 0.011904761904761904), ('Auxiliary verbs', 0.011904761904761904), ('Total pronouns', 0.011904761904761904), ('Space', 0.02380952380952308), ('Past focus', 0.011904761904761904), ('Quantifiers', 0.011904761904761904), ('Conjunctions', 0.023809523809523808), ('Third pers plural', 0.023809523809523808), ('Discrepancy', 0.011904761904761904), ('Negations', 0.011904761904761904), ('Articles', 0.011904761904761904)]
180s Fleece Behind-the-Head Earmuffs	Maybe I‚Äôm just not that sharp but I expected men‚Äôs Bluetooth earmuffs would have speakers in them and connect to my phone. They‚Äôre just earmuffs. They‚Äôre nice and comfortable but I could have gotten just earmuffs for less than half of the price of these.	[('Tentative', 0.022222222222222223), ('Total function words', 0.3070063152760113), ('Negations', 0.022222222222222223), ('Conjunctions', 0.08888888888888889), ('First pers singular', 0.06666666666666667), ('Discrepancy', 0.044444444444444446), ('Auxiliary verbs', 0.044444444444444446), ('Space', 0.022222222222222223), ('Total pronouns', 0.022222222222222223), ('Positive emotion', 0.022222222222222223), ('Causation', 0.022222222222222223), ('Third pers plural', 0.022222222222222223), ('Prepositions', 0.022222222222222223), ('Impersonal pronouns', 0.022222222222222223)]
Echo Plus (2nd Gen) with Philips Hue Bulb	I bought the new echo plus during the black Friday sale this weekend to make use of the Line In feature which I was disappointed that my second generation echo I purchased last year didnt have. I was so excited to see that this changed this year with the new generation... EXCEPT it doest work!! The alexa software that you need to configure it as a 'line in' and not a 'line out' doesnt function, and according to the tech that assisted me, this has been a common problem with the device. I bought it for this reason, as advertised by amazon, and it doest work. The options I were given by the tech was essentially to wait and hope it gets fixed in the future, return it and lose out on the black friday deal if it is fixed in the future and wanted to buy it again at that time once it is, or wait and waste $109 dollars if it isnt fixed in the future once I pass the return window. All for an issue that is 100% AMAZON'S FAULT!! FALSE ADVERTISING and HORRIBLE CUSTOMER SERVICE like I have never experience before from amazon from a problem THEY inflicted and CAN remedy if they choose too! I asked for an extended return window, if they can honor the black friday price if I was to return it and buy it again in the future, and a refund/discount on the product due to it essentially being defective for my purposes, and each request was flat out REJECTED. When I asked to speak to someone else, that wasNOT PERMITTED either. SHAME ON YOU AMAZON, SHAME!!!	[('First pers singular', 0.043478260869565216), ('Total function words', 0.3925471264735436), ('Prepositions', 0.09202457475496 34), ('Interrogatives', 0.007246376811594203), ('First pers plural', 0.001691316598730 8668), ('Space', 0.021739130434782608), ('Articles', 0.010869565217391304), ('Numbers', 0.003623188405797 1015), ('Auxiliary verbs', 0.010869565217391304), ('Conjunctions', 0.065217391304347 82), ('Second person', 0.007246376811594203), ('Negations', 0.007246376811594203), ('Assent', 0.001453907182975311), ('Causation', 0.014492753623188406), ('Cognitive processes', 0.003623188405797 1015), ('Present focus', 0.010869565217391304), ('Past focus', 0.003623188405797 1015), ('Third pers plural', 0.010869565217391304), ('Impersonal pronouns', 0.003623188405797 1015)]
adidas Women's Grand Court Sneaker	I used to love classic adidas shoes but these HURT my heel! The back goes up so high it drills into my skin when I walk.	[('First pers singular', 0.07692307692307693), ('First pers plural', 0.017067063279540388), ('Total function words', 0.2760531864455037), ('Positive emotion', 0.038461538461538464), ('Conjunctions', 0.07692307692307693), ('Impersonal pronouns', 0.038461538461538464), ('Negative emotion', 0.038461538461538464), ('Prepositions', 0.038461538461538464), ('Motion', 0.027951054765315533), ('Space', 0.038461538461538464), ('Interrogatives', 0.038461538461538464)]

Figure S2: Derived scores for 1-Star Reviews.