

Generating Appealing Brand Names

Gaurush Hiranandani¹, Pranav Maneriker¹, and Harsh Jhamtani²

¹ Adobe Research Big Data Experience Lab, Bangalore, India

² Language Technology Institute, Carnegie Mellon University, USA
ghiranan@adobe.com, pmanerik@adobe.com, jharsh@cs.cmu.edu

Abstract. Providing appealing brand names to newly launched products, newly formed companies or for renaming existing companies is highly important as it can play a crucial role in deciding its success or failure. In this work, we propose a computational method to generate appealing brand names based on the description of such entities. We use quantitative scores for readability, pronounceability, memorability and uniqueness of the generated names to rank order them. A set of diverse appealing names is recommended to the user for the brand naming task. Experimental results show that the names generated by our approach are more appealing than names which prior approaches and recruited humans could come up.

1 Introduction

Choosing right brand names for newly launched products, newly formed companies and entities like social media campaigns, apps, websites etc. is critical. In the context of creating brands, it is believed that such a naming decision may well be “the most important marketing decision one can make” [1]. A marketer may often spend a lot of time in coming up with an appealing name which can achieve favorable outcomes on various key performance indicators (KPIs) like website visits, number of customer acquisitions, etc. This becomes critical in scenarios like quickly planned campaigns, where there is not enough time for marketers or authors to come up with an appealing name. However, prior technologies are insufficient to computationally come up with appealing names for such entities based on a provided description. Moreover, rarely is the management provided with interpretable objective criteria upon which a brand name is suggested [1]. This creates a need for an algorithm which automatically generates appealing names from the description of an entity in a justified manner.

This paper has the following contributions. Firstly, we define and infer the importance of various linguistic and statistical features for the task of suggesting names for brands, products or other such entities. Secondly, we propose computational methods to generate brand names given the description of the entity in question. Though coming up with names for entities like brands is considered mostly a creative task, our MTurk based evaluation study determining the appeal of the recommended names shows that names from our method obtain ratings comparable with human-provided names.

Rest of the paper is organized as follows. Section 2 reviews the existing relevant works. Section 3 explains methodology behind generating and ranking names. Section 4 explains the conducted experimental studies and evaluation. In Section 5 we discuss some of the limitations and future work. Lastly in Section 6, we provide conclusions.

2 Related Work

[2] showed characteristics of a ‘good’ name which include short, easy to say, spell, read, understand and easily retrievable from memory. [3] showed consumers use the information they gather from phonemes in brand names to infer product’s attributes. [4] suggested that a recommendation tool for performing the naming task better should aid in ideation as well.

From the perspective of word-generation, there are prior works on password and domain name generation. Some studies have focused on memorability of passwords [5], while tools like *PWGEN* [6] and *Kwyjibo* [7] generate pronounceable passwords and domain names respectively. However, using such tools and being limited to one attribute is not useful for brand-naming generation given the description.

[8] treated blending of words³ as a process to create neologisms. Such blending can be based on various phonetic and syllable alignment techniques, such as those used by [9] and [10]. [11] proposed a computational approach to generate neologisms consisting of homophonic puns and metaphors based on the category and properties of the entity. However, for recommendation purposes, it is important to define a ranking based on appeal of a name given the properties or description of the entity. [11] carry out filtering/ranking during the evaluation by combining the phonetic structure and language model with equal weights. However, ranking based on the appeal of a name is not motivated by the crucial metrics mentioned above. This becomes critical when one needs to recommend a few appealing names rather than generate a large number of names. Further, some online tools⁴ provide names by concatenating random strings and some consulting companies⁵ are engaged in brand naming but do not use any automated processes.

3 Methodology

Figure 3.1 provides an outline of our proposed solution. We do some basic pre-processing on the words in description, followed by expansion of the set of words using an external ontology. We blend the words to generate candidate names.

³ forming a word by combining sounds from two or more distinct words - e.g. *Wikipedia* by blending “Wiki” and “encyclopedia”

⁴ www.online-generator.com, www.namegenerator.biz

⁵ ABC Namebank, A Hundred Monkeys

Then, we score and rank the names based on readability, memorability, pronounceability and uniqueness. Finally we postprocess the ranked list of names to provide a diverse set of suggestions.

As mentioned earlier, we generate candidate names based on blending of syllables present in the word set. The choice of following this approach is based on prior works. [12] provides an annotated dataset of 1000 brand names to understand linguistic creativity involved in naming. In the data, around 20% of the names are created either by juxtaposition or clipping which are morphological mechanisms similar to blending⁶. Further, [8] treats blending of words as a process to create neologisms. This convinced us to generate candidate names based on blending of words.

We prefer using syllables over morphemes as the combination of syllables can generate any names that a combination of morphemes can. The increased number of names generated due to this choice are handled by ranking them and selecting the top candidate names. This is explained in section 3.2.

3.1 Generate Names

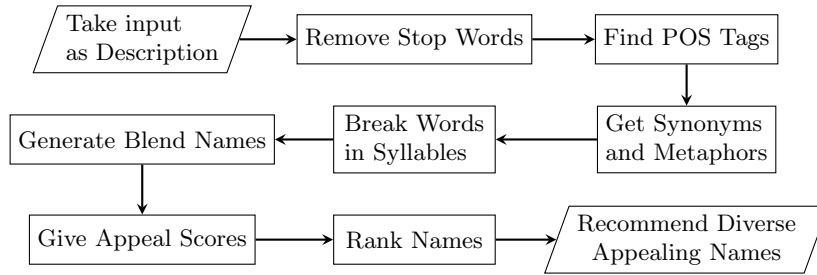


Fig. 1. Overview of the algorithm

The method takes the description of an entity as input. For example, an input can be ‘*Creating an application to split expense wisely.*’ We use words other than stop-words while generating names and call them *root* words. Thereafter, our method looks for part-of-speech (POS) tags [13] of the root words. POS tags are important as words can have a different semantic orientation based on usage and thus have different set of synonyms (used later). Let D denote the set of words along with their POS tags. For example,

$$D = \{(\text{Creating}, \text{Verb}), (\text{Application}, \text{Noun}), (\text{Split}, \text{Verb}), (\text{Expense}, \text{Noun}), (\text{Wisely}, \text{Adverb})\}$$

Next, we use *Wordnet* [14] to obtain synonyms of root words based on POS. Further, we obtain synonyms and metaphors by applying a strategy similar to [11]. We call these words *related words* and attribute the same POS tag

⁶ e.g. *DocuSign* from “Document” and “Signature”

Rule	%	Rule	%
Noun-Adjective	40.10	Adjective-Adverb	3.2
Noun-Verb	8.02	Noun-Noun	36.36
Noun-Adverb	4.81	Verb-Verb	0.00
Verb-Adjective	0.53	Adjective-Adjective	3.28
Verb-Adverb	3.7	Adverb-Adverb	0.00

Table 1. Percentage of Blending Rules

as their root word. Let $C_1, C_2, \dots, A_1, A_2, \dots, S_1, S_2, \dots, E_2, E_2, \dots, W_1, W_2, \dots$ be the related words obtained for *Creating*, *Application*, *Split*, *Expense* and *Wisely* respectively. Further let R denote the list of related words with POS tags. Then,

$$R = \{(Creating, Verb), (C_1, Verb), (C_2, Verb), \dots, (Application, Noun), (A_1, Noun), (A_2, Noun), \dots, (Split, Verb), (S_1, Verb), (S_2, Verb), \dots, (Expense, Noun), (E_1, Noun), (E_2, Noun), \dots, (Wisely, Adverb), (W_1, Adverb), (W_2, Adverb), \dots\}$$

Each word in R is split into syllables using PyHyphen⁷. We attach the same POS tag for each syllable as its parent word. Let us denote the set of syllables of the root and related words along with their POS tags by L . For example,

$$L = \{(Cre, Verb), (At, Verb), (Ing, Verb), (App, Noun), (Li, Noun), (Ca, Noun), (Tion, Noun), \dots\}$$

We observed certain *rules* used in blended names from the description provided by the annotators in [12]. A *rule* is a combination of unordered POS tags of the blended words (syllables). For example, *SplitWise* and *WiseSplit* are created from two syllables whose POS tags are verb and adverb. Hence, they follow the *Verb-Adverb* rule. Table 1 presents the percentages of each rule used in blended names. Note that the percentage of names created using a certain rule are subject to discretionary choice based on the annotator’s bias. As in, for the same entity having similar properties, there can be two different descriptions from two different annotators leading to different empirical estimates. For example, *data platform enables solving problems quickly* or *data platform is quick in problem solving*. Here, we will have the same syllable *quick* in both sentences. However, in the former case, the POS tag is Adverb whereas, in the latter, it is Adjective. For the results shown in Table 1, we have used the POS tags obtained from the description provided by the annotators in [12]. We observed that some of the blending rules were less frequent. Hence, we omitted the rules having less than 1% of the names under their category and called the rest as *allowed rules*. For example, a name using two syllables coming from words having verb as their POS tag is usually not created. An example can be *SplitBreak* formed by the syllables *Split* and *Break*. With respect to the data, this essentially removed three rules but had a significant reduction in the number of names generated.

Finally, the method creates new blended names by joining two or three syllables at a time taken from the permutation set of L , given that the blending is within allowed rules. Let N denote the set of generated names. For the given

⁷ <https://pypi.python.org/pypi/PyHyphen/>

example, some names generated by two syllables are *SplitWise*, *BudSplit* and *BreakOwl*. Note that *break* is a synonym of *split* and *owl* is a metaphor for *wise* obtained from the idiom - *as wise as an owl*. One of the names generated by three syllables is *ExPenseBreak*. We prefer syllables over morphemes as the combination of syllables can generate any name that a combination of morphemes can. The increased number of names generated due to this choice are handled by ranking and selecting the top candidate names. This is explained in the next section.

3.2 Ranking Names

Scores Formulation. Every description can potentially generate thousands of names using the above method. However, it is crucial to rank them for recommendation purposes. Therefore, each name is given a score based on the mathematical formulations of 4 features: readability, pronounceability, memorability, and uniqueness. The scores are normalized in the range of $[0, 1]$ over the English dictionary [15], with 0 and 1 representing the least and the maximum score respectively. Let n and $|n|$ denote a name and its length respectively.

Readability: A good name should be easy to read. Hence, to assign readability score, we use Flesch—Kincaid Reading-Ease Score [16]. Since each name is a single word, readability (denoted by $R(.)$) becomes:

$$R(n) = 205.82 - 84.6 * |\text{syllables}(n)| \quad (1)$$

Where, $|\text{syllables}(n)|$ denotes the number of syllables in n . Later in this section, we will observe formulation in Equation 1 being reduced to number of syllables due to a linear model computing appeal of a name.

Pronounceability: The more permissible the combinations of phonemes is, the more pronounceable the word becomes. We adapt the concept from [17] with some refitting to measure the extent with which a string adheres to the phonotactics of the language. By taking substrings (n-grams) of n of length $l \in \{2, 3, 4\}$ with frequencies from the dictionary [15], we compute certain features as follows:

$$S_l^n = \frac{\sum_{t \in n\text{-grams}(n)} \text{freq}(t)}{|n| - l + 1} \quad (2)$$

Here, $\text{freq}(t)$ is the frequency of the n-gram t in the dictionary. For example,

$$S_2(\text{facebook}) = \frac{fa_{109} + ac_{343} + ce_{438} + eb_{29} + bo_{118} + oo_{114} + ok_{109}}{8 - 2 + 1} = 170.8$$

Feature values for smaller l will be higher than larger l , but feature values for larger l will be more important, since there are around 4000 meaningful 4-letter words in comparison to around 1000 meaningful 3-letter words in English [18]. Hence, the probability of a 4-letter word being meaningfully used will be higher than a 3-letter word. Therefore, weights for S_l^n are assumed as $w_l = \frac{l}{2+3+4}$ for $l = \{2, 3, 4\}$. Finally, we define pronounceability $P(.)$ as:

$$P(n) = w_2 S_2^n + w_3 S_3^n + w_4 S_4^n \quad (3)$$

This formulation is a simple back off model [19], where we always back off to a lower order n-gram with fixed probability.

Memorability: [20] claim that memorable quotes use rare word choices, but at the same time are built upon a scaffolding of common syntactic patterns. Adapting it to blended syllables, we use Meaningful Characters Ratio defined by [17] to capture memorability. It models the ratio of characters of n that comprise a meaningful word. A word is said to be meaningful if it occurs in [18]. We define memorability $M(.)$ as follows:

$$M(n) = \max_{\text{all splits of } n} \frac{\sum_{i=1}^k |s_i|}{|n|} \quad (4)$$

Here, s_i are the meaningful substrings of length ≥ 3 obtained by splitting n and k is their number. For example, If $n = \text{facebook}$,

$$M(\text{facebook}) = \frac{(|\text{face}| + |\text{book}|)}{8} = 1$$

Uniqueness: This feature prefers names having low usage and non-dictionary words. Consider a time series (V, T) for n such that $V = \{v_1, \dots, v_T\}$ and $T = \{t_1, \dots, t_T\}$. t_i s represent consecutive years and v_i represents the normalized usage of n in year t_i as provided by GoogleNgrams [21]. Then, uniqueness $U(.)$ is defined as:

$$U(n) = \frac{\sum_{k=1}^{T_c} v_k * (t_k - t_1)}{\sum_{k=1}^{T_c} (t_k - t_1)} \quad (5)$$

where T_c represents the latest year. The intuition is that less usage in recent years is more important. Further, if GoogleNgrams fails to produce any time series for n , then $U(n)$ is taken to be 1.

Combining Scores. We needed to model the appeal of a brand name depending on the quantitative definitions of the mentioned linguistic features. We asked few annotators to provide descriptions of the entities they want to name. Among the descriptions provided, we randomly chose three of them. Thereafter, we conducted a survey of 20 participants who were shown 3 lists of 15 names generated by our method, one list for each of the three descriptions. They were asked to rank the names in a list from 1 to 15. On an average, the Kendall-Tau correlation between the ‘‘average ranks’’ and the ‘‘individual ranks’’ came out to be 0.66, 0.68 and 0.62 for the 3 lists. High correlation suggested that people give similar rankings of names if shown a description. Hence, we took average ranks to be the ground truth rankings for the 3 lists.

Next we test if the four feature scores described earlier were correlated. In other words, we wanted to test if the individual features provide information not covered by other features. While the pairwise Pearson’s Correlation among

the 4 scores based on 45 names ranged from -0.48 to 0.23 , the Kendall Tau correlation among ranking from each individual score ranged from -0.48 to 0.21 . This implied lack of correlation amongst the scores. Therefore, we defined appeal $A(\cdot)$ of a name to be weighted linear combination of the mentioned scores. That is,

$$A(n) = a_r R(n) + a_p P(n) + a_m M(n) + a_u U(n) \quad (6)$$

Where, $\mathbf{a} = (a_r, a_p, a_m, a_u)$ is the weight vector. Then we applied rank-svm [22] which used the obtained 315 pairwise comparisons to learn the weights showing importance for different features. The learned weights were as follows:

$\mathbf{a} = (2.18, 1.63, 0.91, 1.05)$. Interestingly participants indicated *readability* as the most crucial factor for comparing names amongst the four measures. For $n = SplitWise$, we obtained: $A(n) = 3.71$, $R(n) = 0.77$, $P(n) = 0.04$, $M(n) = 1.0$, $U(n) = 1.0$.

3.3 Recommendation

The user can be recommended with top names as per their appeal scores but since a single syllable may appear in many top names (eg. *fur*, *con* etc. which are meaningful and easy to read), this set of recommendation may not help ideation []. Hence, we diversify the set of names to aid ideation by an update rule [23]. The intuition is that after we choose the top candidate name based on appeal score, we update the appeal scores of the names formed by the same syllables as that of the top candidate name. And then choose the top name from the rest of the names having updated appeal scores. Suppose one chooses n' in the first iteration, then the update for diversity is defined as:

$$A(n) \leftarrow \frac{1}{|m| * |k|} * A(n) \quad \text{for } n \in N \setminus [\{n'\} \cup N'] \quad (7)$$

Here, $|m|$, k , and N' denote the number of common syllables in n and n' , number of syllables in n , and names sharing no common syllable with n respectively. We iteratively choose the best candidate (n') and then update appeal of other names (n) by using Equation 7. The names chosen after some iterations (say 30) are recommended. This ensures that the recommended set of names as a whole become useful for the naming task. For the given example, the top 5 names are *ConTear*, *BreakWise*, *BudSplit*, *BreakOwl* and *DisCleave*.

4 Evaluation

To estimate the quality of the generated names from our method, we conducted an MTurk study and compared our names with two baselines. We created descriptions for 10 entities which usually require brand names. For example, one of the descriptions was *Light-weight software to locate virus on computer*. Our method took 4 minutes on our machine to generate 984 ranked names on an average for a description. We took the top 10 names for each description, names generated from the two baselines (described below) and compared the approaches

through an MTurk study. The code, data, examples and results will be made available online.

4.1 Baselines

Prior Art: [11] describe a method to generate names based on homophonic puns and metaphors by combining natural language processing techniques with various linguistic resources available online. We replicated the work of [11] to generate names. Adapting it our case, the category and properties were provided manually from the descriptions until it output atleast 10 names for the description. After following the original specification, if the number of names generated by this algorithm were fewer than the number required for our experimental setup, we added related properties to the set of properties taken as input. For more details about the approach, see [11]. The output generated from this system was used for further experiments.

Human: 10 participants were recruited to give 10 names for one of the descriptions in 4 minutes (time taken by our method). The participants were given information about the criteria being used for creating new names, i.e. unique and appealing names. This experiment gave 10 human generated names for each description.

4.2 MTurk Survey: Results and Observations

For each description, we created 2 lists of 15 names, each containing 5 names randomly picked from the list of 10 names generated by the three approaches. Table 2 shows a few example inputs to the three approaches and the names generated by them. Then, 100 recruited judges from Amazon Mechanical Turk were shown one of the 20 lists and asked to rate each of the 15 names in it as *Good*, *Fair* or *Bad* based on their relevance to the description and uniqueness. Some of the participants of human experiment in Section 4.1 provided names of currently existing companies. Therefore, the latter instruction was added explicitly to avoid participants rating irrelevant existing names as *Good*. Each list was annotated by 5 judges resulting in 1500 responses.

Table 3 shows percentages of ratings received by names generated by the three approaches considering all 10 descriptions. Our method outperforms the prior art. 16.6% of the names generated by our method received *Good* rating in comparison to 13.2% of the prior art. Similar is the case in *Fair* rating as well. Humans outperform both the automated approaches considering the *Good* rating. However, our method has significantly fewer *Bad* ratings when compared to humans.

We believe that 4 minutes constraint on humans is harsh. They can think of better names if given sufficient time. As a useful observation, one of the participants seemed to be following our approach to generate names using only the root words. The given description was *Showroom of Fabulous Furniture for Decorating Home*. Our method output names like *HomeDec* and *FabFurNi* which were rated mostly as *Fair* whereas the participant generated names like *HomeDecor*

Approach	Input	Output
Our Method	Fabulous furniture to decorate your home	FabFur, MythRate, DressHouse HomeDec, FurDeck
Our Method	Light weight software to locate virus on computer	FeatherTor, PingWare, CleanDen, FaintCate, ClearSet
Prior Art	Category: furniture; Properties: fabulous, decorative, attractive, homely, comfortable	Woodroom, Houly, Floorroom, Dinnel, Bedroose
Prior Art	Category: software; Properties: light, locate, problematic, computable, buggy, safety	Luggyte, Cebuter, Safetyre, Locatr, Coftwarele
Human	Fabulous furniture to decorate your home	Decorature, FabHomes, HomeDecor FabulousHomes, FabFurnish
Human	Light weight software to locate virus on computer	Ubuntu, Nortun, Ad-Blocker, Windows, Web-sites

Table 2. Input and Output by the three approaches

Approach	Good	Fair	Bad
Our Method	16.6%	41.8%	41.6%
Human	20.4%	32.8%	46.8%
Prior Art	13.2%	38.8%	48%

Table 3. Ratings for Generated Names

and *FabFurnish* which were rated mostly as *Good*. This tells us that the method described in this work is indeed a mechanism that humans use to generate names and further, it can also be used for ideation purposes.

Additionally, we calculated nDCG [24] to know whether our method’s rankings match *Good/Fair/Bad* ratings by human judges. In order to define relevance of name n in nDCG formulation, we used 1, 0.5 and 0 as weights for number of *Good*, *Fair* and *Bad* ratings respectively. The nDCG averaged over 10 descriptions was 0.78 indicating that the ranking generated by our method indeed concurs with human rankings.

5 Limitations and Future Work

Our methodology will generate homogeneous names given the same description. Hence, there are opportunities for leveraging enterprise based personalization. Further, we agree that there are brand names like Apple, Fox, etc. which cannot be generated by our approach. However, examples like CarMax, DocuSign, etc. and aforementioned online generators, led us to believe that our approach is one of the ways by which humans create names. In future, we plan to investigate abbreviations, reduplications, and modifications over blended syllables to generate better names.

6 Conclusions

Our work is one of the first approaches to algorithmically generate appealing brand names from description. In addition to being directly used, the recommended names can also aid in ideation. Quantitative definitions of pronounceability, memorability, and uniqueness have been proposed. Further, the set of names generated by us is diverse. The inclusion of diversity aids in the ideation process, providing a rich set of names to any user of the system. Achieving near human results certainly opens the door for automation in this human dominated domain.

Acknowledgments

We thank Dr. Niloy Ganguly for providing valuable comments and feedback.

References

1. Ries, A., Trout, J.: Positioning: The battle for your mind (1981)
2. Robertson, K.: Strategically desirable brand name characteristics. *Journal of Consumer Marketing* **6** (1989) 61–71
3. Yorkston, E., Menon, G.: A sound idea: Phonetic effects of brand names on consumer judgments. *Journal of Consumer Research* **31** (2004) 43–51
4. Little, G., Chilton, L.B., Goldman, M., Miller, R.C.: Exploring iterative and parallel human computation processes. In: *Proceedings of the ACM SIGKDD workshop on human computation*, ACM (2010) 68–76
5. Clements, J.: Generating 56-bit passwords using markov models (and charles dickens). *arXiv preprint arXiv:1502.07786* (2015)
6. Allbery, B.: pwgen-random but pronounceable password generator. *USENET posting in comp. sources. misc* (1988)
7. Crawford, H., Aycock, J.: Kwyjibo: Automatic domain name generation. *Softw. Pract. Exper.* **38** (2008) 1561–1567
8. Bauer, L.: *English word-formation*. Cambridge university press (1983)
9. Kondrak, G.: Phonetic alignment and similarity. *Computers and the Humanities* **37** (2003) 273–291
10. Hedlund, G.J., Maddocks, K., Rose, Y., Wareham, T.: Natural language syllable alignment: From conception to implementation. In: *Proceedings of the Fifteenth Annual Newfoundland Electrical and Computer Engineering Conference*. (2005)
11. Özbal, G., Strapparava, C.: A computational approach to the automation of creative naming. In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*, Association for Computational Linguistics (2012) 703–711
12. Özbal, G., Strapparava, C., Guerini, M.: Brand pitt: A corpus to explore the art of naming. In: *Proceedings of the eighth international conference on Language Resources and Evaluation (LREC-2012)*, Istanbul, Turkey, May, Citeseer (2012)
13. Toutanova, K., Klein, D., Manning, C.D., Singer, Y.: Feature-rich part-of-speech tagging with a cyclic dependency network. In: *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, Association for Computational Linguistics (2003) 173–180

14. Miller, G.A.: Wordnet: a lexical database for english. *Communications of the ACM* **38** (1995) 39–41
15. Goldhahn, D., Eckart, T., Quasthoff, U.: Building large monolingual dictionaries at the leipzig corpora collection: From 100 to 200 languages. In: *LREC*. (2012) 759–765
16. Kincaid, J.P., Jr, R.P.F., Rogers, R.L., Chissom, B.S.: Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. Technical report, DTIC Document (1975)
17. Schiavoni, S., Maggi, F., Cavallaro, L., Zanero, S.: Phoenix: Dga-based botnet tracking and intelligence. In: *Detection of intrusions and malware, and vulnerability assessment*. Springer (2014) 192–211
18. Webster, M.: The Official Scrabble Players Dictionary. “Springfield” (2005)
19. Chen, S.F., Goodman, J.: An empirical study of smoothing techniques for language modeling. In: *Proceedings of the 34th annual meeting on Association for Computational Linguistics, Association for Computational Linguistics* (1996) 310–318
20. Danescu-Niculescu-Mizil, C., Cheng, J., Kleinberg, J., Lee, L.: You had me at hello: How phrasing affects memorability. In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1, Association for Computational Linguistics* (2012) 892–901
21. Michel, J.B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Pickett, J.P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., et al.: Quantitative analysis of culture using millions of digitized books. *science* **331** (2011) 176–182
22. Joachims, T.: Optimizing search engines using clickthrough data. In: *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM* (2002) 133–142
23. Carbonell, J., Goldstein, J.: The use of mmr, diversity-based reranking for re-ordering documents and producing summaries. In: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval, ACM* (1998) 335–336
24. Järvelin, K., Kekäläinen, J.: Cumulated gain-based evaluation of ir techniques. *ACM Transactions on Information Systems (TOIS)* **20** (2002) 422–446