Putting to test the Affective-Aesthetic Potential

Boot, Peter

peter.boot@huygens.knaw.nl
Huygens Institute for the History and Culture of the Netherlands

Koolen, Marijn

marijn.koolen@di.huc.knaw.nl
Huygens Institute for the History and Culture of the
Netherlands; DHLab, KNAW Humanities Cluster

Mussman, Ole

o.mussmann@esciencecenter.nl
Netherlands eScience Center

Schnober, Carsten

c.schnober@esciencecenter.nl
Netherlands eScience Center

van Hage, Willem

w.vanhage@esciencecenter.nl
Netherlands eScience Center

van Zundert, Joris

joris.van.zundert@huygens.knaw.nl
Huygens Institute for the History and Culture of the
Netherlands; DHLab, KNAW Humanities Cluster

Introduction

We investigate the Affective-Aesthetic Potential (AAP), a psycholinguistic word measure meant to represent the beauty of words (Jacobs 2017). Jacobs and Kinder (2020) show that the average AAP value predicts how well readers like passages in novels. In this paper, we investigate the AAP's potential in the context of a corpus ca. 11.000 Dutch-language fiction books for which we have text and sometimes reviews and ratings.

The AAP was introduced by Jacobs (2017). At the basis of its computation are two lists of 62 seed words (positive and negative). The AP represents the semantic distance to the seed words. Jacobs presents the seed words as 'tentative' but apparently, they have not been revised since. The SentiArt framework (Jacobs 2019, Jacobs and Kinder 2020) expands on this approach. Here, a vector space model is computed using word2vec. From seed words, values are computed for various (approximations of) psycholinguistic variables. We focus on valence and AAP.

We present some computational experiments around valence and AAP in the context of our corpus. All reported results are tentative; we will recompute them before the conference with a larger corpus of books and reviews.

Computation of AAP and Valence ratings

We compute SentiArt valence and AAP values as follows:

- We use word2vec to create word embeddings from the complete text of our novels (931M words).
- We translate the AAP seed words into Dutch.
- We compute AAP values following Jacobs and Kinder (2020).

Our data are available at: https://github.com/impact-and-fiction/putting-to-test-aap.

SentiArt-predicted vs. empirical values of valence

Jacobs gives valence and AAP values for Dutch words. We test SentiArt's usability by comparing Jacobs' valences with those established empirically by Moors et al. (2013) and with the words in the LiLaH sentiment lexicon (Ljubešić et al. 2020).
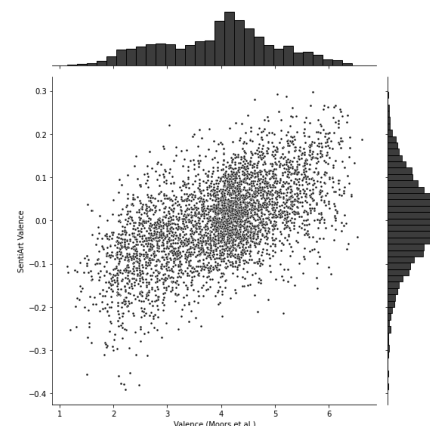

Figure 1. The SentiArt valence and Moors et al. valence scores of individual words.

Figure 1 presents the SentiArt valences with corresponding scores from Moors et al. The correlation is substantial (0.70). However, there are many words with a low score on one scale but high on the other. The results on the LiLaH lexicon (Figure 2, not discussed here) are similar. Further study is needed to find out what this might mean in application scenarios.
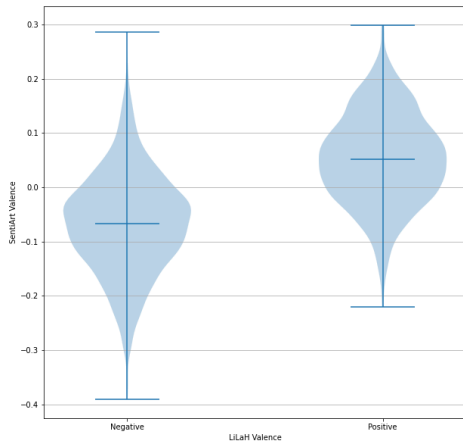
Figure 2. Distribution of SentiArt scores for 4490 words in the LiLaH dataset, differentiated by their binary valence in LiLaH.

AAP values based on Wikipedia corpus vs. values based on literary corpus

Jacob's Dutch AAP values were computed on Wikipedia word embeddings. We use word embeddings based on our novel corpus. Figure 3 shows that the embeddings based on novels give a larger amplitude of scores than those based on Wikipedia, suggesting that in novels, much of the vocabulary is more associated with emotion words than in Wikipedia. This provides a good reason to work with the book embeddings-based variables in our further analyses.
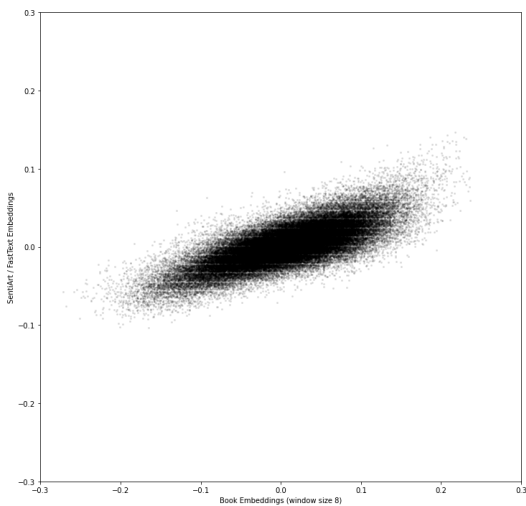

Figure 3. SentiArt AAP scores based Wikipedia (Y-axis) and on our novels.

Relation of AAP to valence over the course of a single work of fiction

Many of the AAP seed words are clearly related to valence (e.g. good, bad). It is to be expected that valence and AAP values are closely aligned. To test this, we plot both over the course of one novel (Figure 4). It appears both constructs are very similar indeed (correlation 0.85). We will look further into the locations where the AAP and valence patterns differ.
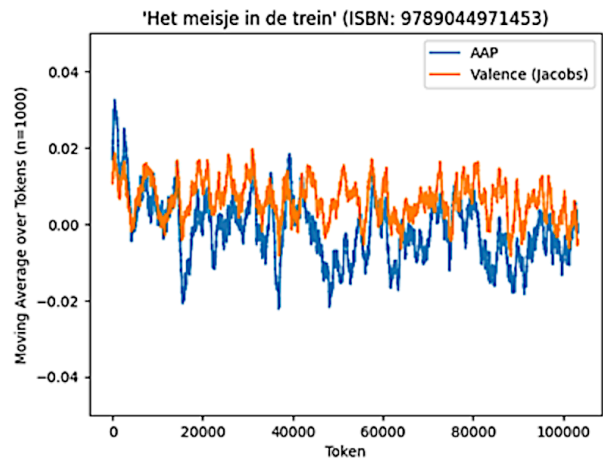

Figure 4. Moving averages of AAP and (SentiArt-based) valence.

Correlation between book level AAP and user ratings

Jacobs and Kinder (2020) report that AAP values for segments of a long story predict liking by readers. We investigate whether this relation still exists at the book level, by computing the correlation between book-level averaged AAP and readers' ratings, for books for which we have both text and reviews (670 books). We use the 'Lens' method described by Jacobs and Kinder.

The correlation is -.11, and therefore weakly negative, contrary to our expectation. Figure 5 plots the values.
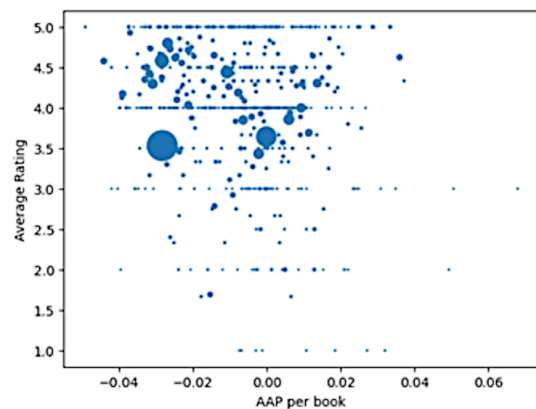
Figure 5. Average rating vs. average AAP value per book. The size of each dot represents the number of reviews.

There are a number of potential explanations for the absence of a relationship. Are only the last sections of the book important? Are ratings cognitive rather than affective? Should we control for genre? Should we look at fractal-based patterns (Gao et al. 2016)? We will come back to this at the conference.

Correlation between book level AAP and stylistic impact

As AAP should be particularly sensitive to the aesthetic component, we compute the correlation between book-level averaged AAP and the stylistic impact expressed in user reviews for that book (Boot and Koolen 2020).
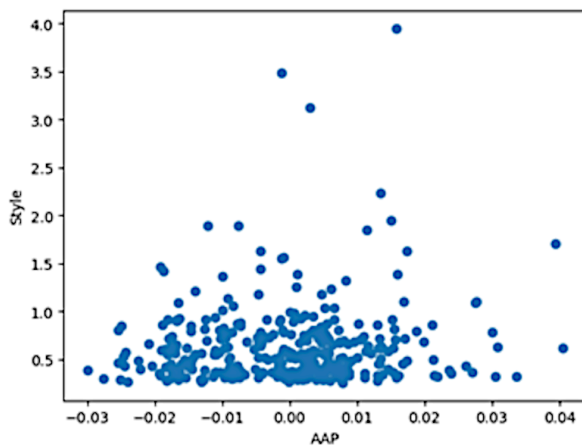

Figure 6. Average stylistic impact vs. average AAP value per book (n=670).

We plot the result in Figure 6. We find no significant correlation (r = 0.08, p = 0.14). We look for explanations based on a.o. the suggestions above.

Discussion and conclusion

Based on our results up to now we can say that the AAP represents an interesting approach to the study of literary response, but more work is needed. Apart from what we already noted above this includes:

- As indicated by Jacobs & Kinder (2020), the AAP methodology is not designed for long texts. We should look deeper into what AAP can do for longer text.
- AAP values are based on lists of positive and negative seed words. Does it make sense that words with a large negative AAP value neutralize ones with large positive values? Do positive and negative affects cancel out each other in literary texts? This is especially relevant given the somewhat arbitrary definition of AAP in Jacobs (2017).

- Following Jacobs, we have used only words in the top and bottom percentiles of the AAP value range (the "Lens" method). However, the impact of this choice needs to be investigated both empirically and analytically.

Bibliography

Boot, Peter (2017): "A Database of Online Book Response and the Nature of the Literary Thriller", in: In Conference Proceedings *DH2017*.

Boot, Peter / **Koolen, Marijn** (2020): "Captivating, splendid or instructive?: Assessing the impact of reading in online book reviews", in: *Scientific Study of Literature* , *10,* 1, 35-63.

Gao, J., et al. (2016). "A multiscale theory for the dynamical evolution of sentiment in novels", in: *2016 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC)*, IEEE.

Jacobs, Arthur M. (2017): "Quantifying the beauty of words: a neurocognitive poetics perspective", in: *Frontiers in Human Neuroscience*, *11*, 622.

Jacobs, Arthur M. (2019): "Sentiment analysis for words and fiction characters from the perspective of computational (Neuro-) poetics", in: *Frontiers in Robotics and AI* 6. DOI: /10.3389/frobt.2019.00053

Jacobs, Arthur M. / **Kinder, Annette** (2020): "Computing the affective-aesthetic potential of literary texts", in: *AI* 1, 1: 11-27.

Koolen, Marijn / **Boot, Peter** / **Van Zundert, Joris J.** (2020): "Online Book Reviews and the Computational Modelling of Reading Impact", in: *Proceedings http://ceur-ws.org*.

Ljubešić, Nikola / **Markov, Ilia** / **Fišer, Darja** / **Daelemans, Walter** (2020): "The lilah emotion lexicon of croatian, dutch and Slovene", in: In *Proceedings of the Third Workshop on Computational Modeling of People's Opinions, Personality, and Emotion's in Social Media* : 153-157.

Moors, Agnes / **De Houwer, Jan** / **Hermans, Dirk** / **Wanmaker, Sabine** / **Van Schie, Kevin** / **Van Harmelen, Anne Laura** / **...** / **Brysbaert, Marc** (2013): "Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words", in: *Behavior research methods 45,* 1: 169-177.