A computational model of poetry appreciation based on a spreading activation network and the incongruity resolution theory

Chota Kameya (c.kameya@uec.ac.jp)
Tomoki Miyamoto (miyamoto@uec.ac.jp)
Akira Utsumi (utsumi@uec.ac.jp)

Department of Informatics, The University of Electro-Communications
1-5-1, Chofugaoka, Chofushi, Tokyo 182-8585, Japan

Abstract

In the studies of aesthetic appreciation, poetry has received less attention than other art forms. Furthermore, although a number of recent studies have focused on the factors that influence poetry appreciation, very few attempts have been made to explain the appreciation process. In this study, to explore the cognitive process of poetry appreciation, we propose a computational model of poetry appreciation based on the incongruity resolution theory, which has been proposed to explain the process of evoking poetic effects. Assuming that poetry reading comprehension is a process of evoking mental imagery from words in a poem, we model it as spreading activation in a semantic network. The notion of incongruity resolution is then modeled as the difference of the diversity of activation in the network. The computational model was tested using multiple regression analysis with simulation results as independent variables and human ratings for aesthetic appeal as a dependent variable. The analysis demonstrated that the degree of incongruity resolution computed by the model accounted for a significant portion of the variance in aesthetic appeal. This finding provides evidence for the validity of the model and suggests its potential for a more comprehensive model of poetic appreciation in literary works.

Keywords: poetry; aesthetic appreciation; spreading activation network; incongruity resolution theory; computational model

Introduction

Aesthetic appreciation is a process in which an observer enjoys the aesthetic value evoked by works of art. It is a complex process that is influenced not only by the characteristics of art works, but also by the subjective perception of the observers. Therefore, aesthetic appreciation has attracted the interest of researchers in many disciplines, such as psychology, cognitive science, linguistics, and neuroscience.

Studies of the aesthetic appreciation of artworks have often focused on each domain of the artwork because it is likely that the processes and factors influencing them differ according to the domain of the artwork (Mehl, Gugliano, & Belfi, 2023). For example, a comprehensive model of the process of aesthetic appreciation has been proposed for the visual arts in neuroscience (e.g., Chatterjee & Vartanian, 2016) and information processing (e.g., Leder, Belke, Oeberst, & Augustin, 2004; Leder & Nadal, 2014). However, little attention has been paid to the aesthetic appreciation of poetry, although poetry has long been a popular art form.

Recently, an increasing number of studies have addressed the aesthetic appreciation of poetry. These studies can be divided into two approaches. The first one is a stimulus-oriented approach, which focuses on the characteristics of poetry that influence aesthetic appreciation or evoke aesthetic effects. The existing studies of this approach have shown that the aesthetic appreciation of poetry is influenced by a variety of properties of poetry, such as meter and rhyme (Obermeier et al., 2013), semantic cohesion (Hugentobler & Lüdtke, 2021), and title type (Maruyama & Ishizu, 2025). The second approach is a reader-oriented approach, which focuses on the mental state and subjective evaluation of the reader. Of particular interest is the finding that the vividness of the mental imagery evoked by poetry reading influences aesthetic evaluation (e.g., Belfi, Vessel, & Starr, 2018; Hitsuwari & Nomura, 2021).

However, many of these studies do not address the cognitive processes involved in aesthetic appreciation; they only examine the factors that influence aesthetic appreciation of poetry. Some stimulus-oriented studies (Obermeier et al., 2013; Delmonte, 2021; Maruyama & Ishizu, 2025) have discussed the relationship between the characteristics of poetry and the cognitive (or processing) fluency theory that is relevant to aesthetic appreciation. Some neuroscience studies (Jacobs, 2015; Obermeier et al., 2016) have attempted to explain and understand the process of poetry appreciation at the neurological level. However, there are no studies that can explain how poetic or aesthetic effects are evoked through the process of reading comprehension.

In this paper, therefore, we propose a computational model of poetry appreciation as a first step toward a comprehensive understanding of the cognitive mechanisms underlying poetry appreciation. The model is based on the incongruity resolution theory, which posits that poetic effects are evoked when the imposed incongruity is resolved during the subsequent stages of comprehension. To model the process of poetry comprehension during which incongruity resolution occurs, we assume that poetry reading comprehension is a process of experiencing mental images evoked by the words in a poem, and model this process as spreading activation on a semantic network. Incongruity resolution is then modeled using the difference in diversity of word activation in the network between the adjacent lines of the poem. To test the validity of the proposed model, we conducted an evaluation experiment in which human ratings of the aesthetic appeal of poems were regressed on their degree of incongruity resolution computed by the model. The proposed model is fully computational in that it takes a poem as input and computes the degree of incongruity resolution as a predictor value for aesthetic appeal.

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

This model makes a novel contribution to the research on poetry appreciation, because there have been no studies that explain aesthetic appreciation directly through the process of poetry reading comprehension.

Incongruity Resolution Theory

The incongruity resolution theory (Utsumi, 2002, 2006) is one of the theories that explains the process of evoking poetic effects. According to this theory, poetic effects are evoked by the following process.

1. The incongruity of linguistic expressions imposes a cognitive load on the reader.
2. Poetic effects are evoked when the reader resolves the incongruity by generating a sufficiently rich interpretation.

This theory predicts that the greater the cognitive load and the more successfully the incongruity is resolved, the greater the poetic effects (Utsumi, 2002). Incongruity resolution was originally proposed as a model to explain the process of humor appreciation (Suls, 1972). It is also consistent with the relevance-theoretic view of poetic effects (Pilkington, 2014).

The psychological plausibility of the incongruity resolution theory is demonstrated in the case of metaphor appreciation. Utsumi (2005) assumed the semantic dissimilarity between the vehicle and tenor of a metaphor to be the imposed incongruity, and examined its relationship with the diversity of interpretation and poeticality of the metaphor. He found that higher dissimilarity leads to greater interpretive diversity, and greater interpretive diversity leads to higher poeticality ratings for comprehensible metaphors. This result lends support to the validity of the incongruity resolution theory.

Model

In this paper, poetry reading comprehension is conceptualized as the process of evoking mental images from the words in a poem, which is modeled as a spreading activation in a semantic network. Furthermore, we model cognitive load and incongruity resolution using the diversity of word activation across the entire semantic network.

Modeling the Process of Poetry Reading Comprehension

A number of aesthetic studies have argued that mental imagery plays a crucial role in the reading comprehension of poetry. People experience visual (or sometimes auditory) mental images evoked by reading poetry. These mental images not only evoke aesthetic or poetic effects, but also influence the comprehension and appreciation of the content (Liu, 2022). Several studies (e.g., Belfi et al., 2018; Hitsuwari & Nomura, 2021) have shown that the more vivid the images evoked by poetry, the greater the aesthetic value perceived by readers. Furthermore, it has been reported that the vividness of imagery is a crucial factor in the aesthetic evaluation of poetry, especially when compared to painting or music (Mehl et al., 2023).

According to the findings of these studies, poetry comprehension can be conceptualized as the process of forming a mental image of the poetry. Furthermore, the comprehension of a poem depends on the compositional semantics constructed from the elements of the poem (Johnson-Laird & Oatley, 2022). Given the conception of poetry as a collection of words, the mental imagery of poetry is formed by the images evoked by the individual words. Therefore, in this paper, we reasonably assume that poetry reading comprehension is a process of evoking images from the sequence of words in a poem.

The process of evoking imagery in the model In this paper, the process of evoking mental imagery from the sequence of words in a poem is modeled as spreading activation in a semantic network (Collins & Loftus, 1975). The spreading activation model is a framework that mimics semantic processing, particularly semantic priming. A semantic network consists of nodes representing words and links representing the relationship between words. Semantic processing in the spreading activation model is modeled as a process where the activation of words is spread to other words through links. A number of studies (e.g., Vitevitch, Ercal, & Adagarla, 2011; Rotaru, Vigliocco, & Frank, 2018; Siew, 2019) have proposed a computational formalization of spreading activation to explain the process of lexical access or semantic priming. In this paper, we use a similar formalization to these studies. Note that the proposed model is consistent with the finding that semantic priming influences aesthetic appreciation (Faerber, Leder, Gerger, & Carbon, 2010); this suggests the plausibility of our approach of using the spreading activation model.

In this paper, the degree of activation $a_i(t)$ $(0.0 \leq a_i(t) \leq 1.0)$ of a word i ($1 \leq i \leq |V|$ where V is the vocabulary of the semantic network) at step t is defined by the following equations.

$$a_i(t) = \tanh\{r \cdot a_i(t-1) + (1-r) \cdot fam_i \cdot \sum_j w_{ji} a_j(t-1)\} \quad (1)$$

$$w_{ij} = \frac{similarity(i,j)}{\sum_k similarity(i,k)} \quad (2)$$

$$similarity(i,j) = \begin{cases} cos(i,j) & cos(i,j) \geq \theta \\ 0.0 & \text{otherwise} \end{cases} \quad (3)$$

In Equation (1), r $(0.0 < r < 1.0)$ denotes the proportion of activation retained for the next step, fam_i $(0.0 \leq fam_i \leq 1.0)$ denotes the familiarity score of the word i, w_{ji} denotes the weight of the link from the word j to the word i, and $\tanh(\cdot)$ is the hyperbolic tangent function. In Equation (3), $cos(i,j)$ denotes the cosine similarity between the words i and j, and θ $(0.0 \leq \theta < 1.0)$ denotes the threshold of cosine similarity used to determine whether a link is established between words.

Reading comprehension process in the model To model the process of poetry reading comprehension, we extend the spreading activation model so that multiple words are activated sequentially in the order in which they appear in a poem.

Input: A poem $P = l_0, l_1, ..., l_N$ where $l_j = w_{j1}, ..., w_{jn_j}$
Output: $CL(P)$ and $IR(P)$

1. For l_j in $l_0, l_1, ..., l_N$:
   (1a) For w_{jk} in $w_{j1}, ..., w_{jn_j}$:
      i. Activate the word w_{jk} by setting its degree of activation $a_i(0)$ to 1.0.
      ii. For $t = 1$ to T:
         - Spread the activation across the semantic network, that is, compute the degree of activation $a_i(t)$ of all words by Equation (1).
   (1b) Compute $H(l_j)$ by Equation (4).
2. Compute $CL(P)$ and $IR(P)$ by Equations (7) and (8).

Figure 1: An overall algorithm of the proposed model for poetry appreciation

The process of reading comprehension in the proposed model is shown in Figure 1. Let a poem P be a sequence of the title l_0 and N lines, that is, $P = l_0, l_1, ..., l_N$, and each line l_j be a sequence of n_j content words, that is, $l_j = w_{j1}, ..., w_{jn_j}$. In the proposed model, reading comprehension proceeds line by line. Each line l_j is comprehended by repeating the spreading activation process described above, that is, the process of evoking mental images from the word w_{jk} to be read.

Modeling the Incongruity Resolution Theory

The notion of incongruity (or cognitive load) and its resolution can be quantified from the activation state of the entire semantic network. Incongruity occurs when mental images evoked after reading a line l_j greatly differ from those evoked by its preceding line l_{j-1}. In this case, many words that were not activated after reading l_{j-1} are likely to be activated by the line l_i, and thus the diversity of network activation tends to increase. In contrast, incongruity is resolved (or cognitive load is reduced) when mental images evoked by a line l_j are similar to those evoked by the preceding line l_{j-1}, that is, they meet the expectations just after reading the preceding line. In this case, many words activated by the preceding line are likely to become more activated, and thus the diversity of network activation decreases.

Given the above discussion, we model the cognitive load and incongruity resolution as the difference of diversity between adjacent lines. The diversity of network activation after reading the line l_j is defined as Shannon entropy $H(l_j)$.

$$H(l_j) = -\sum_{i=1}^{|V|} p_i(T) \log p_i(T) \tag{4}$$

$$p_i(T) = \frac{a_i(T)}{\sum_{k=1}^{|V|} a_k(T)} \tag{5}$$

When the entropy difference $\Delta H(l_j)$ between lines l_j and l_{j-1}, defined by Equation (6), is positive, it implies that incongruity occurs (or cognitive load is imposed) after reading the line l_j. The negative difference of entropy implies that incongruity is resolved (or cognitive load is reduced). Therefore, the degree of cognitive load $CL(P)$ for a poem P is defined as the average positive entropy difference by Equation (7). The degree of incongruity resolution $IR(P)$ for a poem P is defined as the absolute value of the average negative entropy difference computed by Equation (8).

$$\Delta H(l_j) = H(l_j) - H(l_{j-1}) \tag{6}$$

$$CL(P) = \frac{1}{N} \sum_{j=2}^{N} \Delta H(l_j) \mathbb{1}_{(\Delta H(l_j)>0)} \tag{7}$$

$$IR(P) = \frac{1}{N} \sum_{j=2}^{N} |\Delta H(l_j)| \mathbb{1}_{(\Delta H(l_j)<0)} \tag{8}$$

In these equations, $\mathbb{1}_{(cond)}$ is an indicator function that returns 1 if the condition $cond$ is true and otherwise 0. Note that the entropy difference $\Delta H(l_1)$ between the title and the first line is excluded in averaging the entropy differences in Equations (7) and (8).

Model Testing

To test the validity of the model, we collected human ratings of aesthetic value for 20 poems and examined the extent to which these ratings can be explained by the cognitive load $CL(P)$ and incongruity resolution $IR(P)$ computed by the proposed model for the same poems.

Materials

Twenty Japanese poems (10 lyric and 10 epic ones) published in the 20th and 21st centuries were chosen for model testing. The average number of lines in these poems was 13.05, and the standard deviation was 4.19.

Questionnaire Survey

We conducted a questionnaire survey to collect human ratings of aesthetic value. This survey was approved by the Ethics Committee of the University of Electro-Communications.

Aesthetic dimension To measure the aesthetic value of poems, we used aesthetic appeal ("I find this poem aesthetically appealing") as an aesthetic dimension for rating, which has been used in other studies on poetry appreciation (Belfi et al., 2018; Hitsuwari & Nomura, 2021). We also used the following three dimensions: imagery vividness I ("It is easy to visualize the imagery evoked by this poem"), imagery vividness II ("The imagery evoked by this poem is vivid"), liking ("I like this poem"). These dimensions were used for additional analyses we do not discuss in this paper. Some of them were used as explanatory variables for aesthetic appeal (Belfi et al., 2018; Hitsuwari & Nomura, 2021). All dimensions were rated on a 7-point Likert scale (1: strongly disagree, 7: strongly agree).

Procedure Participants were recruited through a crowdsourcing service and participated in the survey via a web application. After reading an instruction of the survey, they were asked to read and rate one poem at a time. The presentation order of the poems was randomized for each participant. Each poem was displayed one line at a time,and participants pressed the button to read the next line. After reading each poem, they rated it on the four dimensions described above. After rating a poem, they were asked to indicate whether they had ever read it. For each poem, 20 participants rated using the above procedure.

Model Simulation

Data For the vocabulary V of the semantic network, we used a total of 49,613 content words included in the NTT database "Lexical Properties of Japanese" or the 20 poems. Proper nouns and words that do not have a word vector were excluded. All content words in the database and poems were lemmatized. For the computation of cosine similarity $cos(i, j)$ between words in Equation (3), we used the word vectors of 200 dimensions from Japanese Wikipedia entity vectors [1]. For the familiarity score fam_i of Equation (1), we used the data on word familiarity included in the NTT database. Because the original data ranges from 1 to 7, we converted it to a range of 0 to 1. Words that appeared in the poem but were not included in the database were assumed to have a familiarity score of 1.

Procedure Table 1 shows model parameters: θ, a threshold of links that determines the density of the network; r, the proportion of retained activation; and T, the activation time. First, five semantic networks were created according to different values of θ. Table 2 shows the density of each semantic network. When the value of θ is 0.0 or 0.1, the density is extremely high and the network is close to a complete graph. When the value of θ is 0.3 or 0.4, the density is very low and the network is quite sparse. In contrast, the network with θ of 0.2 is moderately dense between the two extremes.

Subsequently, for each of the all combinations of parameter values listed in Table1, we computed the degree of cognitive load $CL(P)$ and incongruity resolution $IR(P)$ for the 20 poems using the model shown in Figure 1. As a result, 100 ($= 5 \times 4 \times 5$) values of $CL(P)$ and $IR(P)$ were obtained for each poem.

Evaluation

To verify the validity of the model, we examined whether and to what extent the simulation result (i.e., the degree of cognitive load and incongruity resolution) can explain the human ratings obtained in the questionnaire survey. To this end, we conducted a multiple regression analysis for each of the 100 simulation results, with the cognitive load $CL(P)$ and the incongruity resolution $IR(P)$ as independent variables and the average ratings of aesthetic appeal as a dependent variable.

[1] https://www.cl.ecei.tohoku.ac.jp/~m-suzuki/jawiki_vector/

Table 1: Parameter values used for the simulation

Parameter	Value
θ	0.0, 0.1, 0.2, 0.3, 0.4
r	0.1, 0.2, 0.3, 0.4
T	1 ~ 5

Table 2: Density of each semantic network

θ	Number of links	Density of semantic network
0.0	2,456,444,854	0.998
0.1	2,274,434,290	0.924
0.2	1,484,862,728	0.603
0.3	643,232,252	0.261
0.4	204,512,958	0.083

Results

Any rating data from the participants who responded that they had ever read the poem were excluded from the analysis, because their prior experience with the poem might have influenced their aesthetic judgment. In addition, one poem was excluded from the analysis as an outlier because its level of incongruity resolution was extremely high.

Table 3 lists the results of the multiple regression analyses with the top five highest coefficients of determination R^2. The high coefficients of determination $R^2 = 0.587$ and its large effect size (Cohen's $f^2 = 1.42$) indicate that the proposed model can considerably explain the aesthetic appeal of poetry. Table 3 also shows that the degree of incongruity resolution $IR(P)$ accounted for a significant proportion of variance in aesthetic appeal, but the degree of cognitive load $CL(P)$ did not. In addition, the parameter θ, which determines network density, was 0.2 in all the top five results.

Figure 2 shows the scatter plots between aesthetic appeal and two variables computed by the model with the parameter values $r = 0.1$, $T = 1$, and $\theta = 0.2$ at which the highest R^2 was achieved. As depicted in Figure 2, there was a highly positive correlation between aesthetic appeal and incongruity resolution ($r = .766$), but cognitive load was not correlated with aesthetic appeal.

Table 3: Results of the multiple regression analysis with the top five highest coefficient of determination

				b^*	
r	T	θ	R^2	Cognitive load	Incongruity resolution
0.1	1	0.2	0.587	0.021	0.773***
0.3	4	0.2	0.543	-0.297	0.883***
0.2	3	0.2	0.538	-0.174	0.829**
0.1	3	0.2	0.537	-0.132	0.810**
0.1	5	0.2	0.529	-0.134	0.802**

* $p < .05$, ** $p < .01$, *** $p < .001$

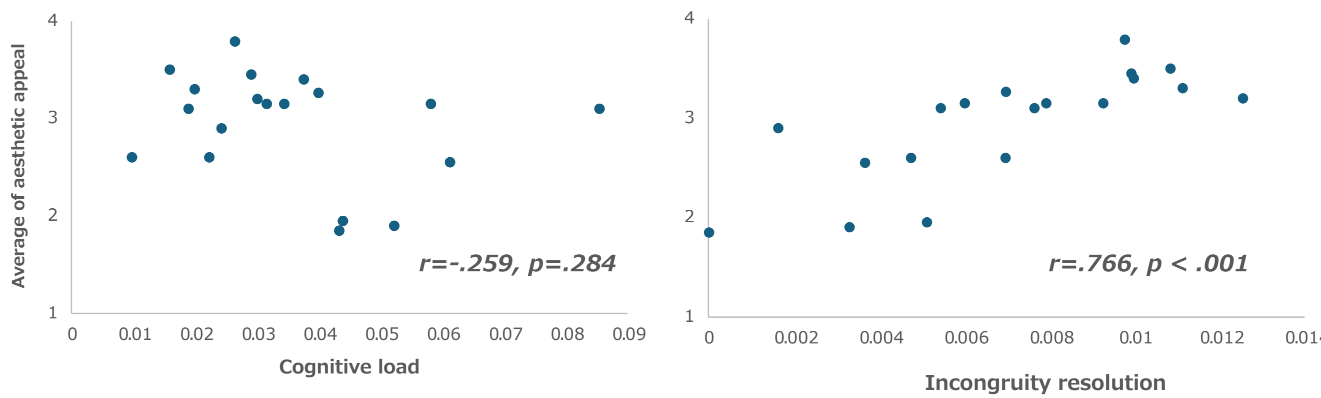

Figure 2: Scatter plot between aesthetic appeal and two independent variables at the highest R^2 ($r = 0.1, T = 1, \theta = 0.2$).

Discussion

The large effect size of the multiple regression analysis can be interpreted as evidence supporting the validity of the proposed model. The finding of a significant positive correlation between the incongruity resolution $IR(P)$ and aesthetic appeal is consistent with the incongruity resolution theory, suggesting the validity of the theory.

However, the absence of a significant correlation between the degree of cognitive load $CL(P)$ and aesthetic appeal may cast doubt on the validity of the proposed model. One possible explanation is that the incongruity resolution theory does not necessarily claim that the imposition of a cognitive load (or the occurrence of an incongruity) is a sufficient condition for evoking aesthetic appeal. If this claim is correct, then the obtained result is fully consistent with the incongruity resolution theory. One difficulty with this explanation, however, is that the definition of the degree of incongruity resolution $IR(P)$ as the difference in entropy between adjacent lines does not presuppose the occurrence of an incongruity. It would be interesting for further work to develop a variable that better reflects the incongruity resolution theory.

Another possible explanation for the lack of correlation with the cognitive load is that the proposed model is not valid as a model of poetry appreciation. The basic tenet of the proposed model is that the evocation of mental images from words can be modeled as spreading activation in a semantic network, and that the incongruity resolution theory can be explained as differences in the diversity of network activation. If this tenet is correct, the relationship among cognitive load, incongruity resolution, and aesthetic judgment would remain unchanged even if non-basic parts of the model are changed. In other words, if some variant models yield results that are consistent with the original, this suggests that cognitive load may not affect aesthetic judgment.

Therefore, we conducted the same evaluation experiment using the following two variant models: (a) the model in which the hyperbolic tangent function is removed from Equation (1) and thus $a_i(t) \geq 0$ and (b) the model using a different semantic network with a different vocabulary, i.e., a set of 15,000 most frequent words on entity vectors. Both models yielded an almost identical result. For the model (a), the five highest coefficients of determination ranged from .529 to .587, and aesthetic appeal was significantly influenced by incongruity resolution ($b^* = .774 \sim .885, p < .01$), but not by cognitive load ($b^* = -.303 \sim .022$). For the model (b), the top five coefficients of determination parameters ranged from .381 to .439, and aesthetic appeal was significantly influenced by incongruity resolution ($b^* = .574 \sim .664, p < .05$), but not by cognitive load ($b^* = .004 \sim .120$). These consistent results indicate that cognitive load alone may not be a sufficient condition for aesthetic judgment.

The result that the same parameter value of θ achieved the top five results listed in Table 3 suggests that the density of the semantic network is an important component for a plausible model of poetry appreciation. As shown in Table 2, the density of the semantic network with $\theta = 0.2$ is moderate and differs significantly from other networks. Hence, a finer setting of θ may lead to a better performance of the model.

The poetry appreciation model proposed in this paper has some limitations. An important limitation is that the proposed model cannot explain the influence of phonological or rhythmic features of poetry such as meter and rhyme on aesthetic appreciation. As mentioned in the section of introduction, it has been found that these phonological features have a significant impact on aesthetic appreciation of poetry (e.g., Obermeier et al., 2013). The appreciation of the poem excluded from the analysis as an outlier may be explained by these phonological features. For a more comprehensive explanation of the aesthetic value of poetry, these features must be integrated into the proposed computational model. A further limitation of the model is its exclusive reliance on lexical semantic processing, excluding any other semantic processes such as sentence comprehension and figurative understanding. The proposed model is a simplified model of the reading comprehension process based on lexical processing, and thus it ignores the grammatical structure of lines and figurative expressions. In particular, a figurative expression (or a figure of speech) is an important device that poets often use in their poems. Furthermore, this paper focuses only on the role of mental imagery and aesthetic appeal in the appreciation of poetry. However, other factors, such as emotions, play an important role in the aesthetic experience, and the model in this paper cannot explain them. Further research is needed to overcome these limitations.

Conclusion

In this paper, we developed the model to explain the cognitive process of poetry appreciation. Assuming that the mental imagery of a poem is represented as a collection of associated words, we modeled the process of poetry reading comprehension as spreading activation in a semantic network. The notion of incongruity resolution was quantified as differences in the diversity of word activation between the adjacent lines of a poem. The multiple regression analysis demonstrated that incongruity resolution significantly affected aesthetic appeal and the proposed model showed high explanatory power. The computational model proposed in this paper is expected to contribute to the development of cognitive science research on poetry appreciation and to serve as a basis for future research.

References

Belfi, A. M., Vessel, E. A., & Starr, G. G. (2018). Individual ratings of vividness predict aesthetic appeal in poetry. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(3), 341–350.

Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, *1369*(1), 172–194.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428.

Delmonte, R. (2021). Cognitive models of poetry reading. In M. Danesi (Ed.), *Handbook of cognitive mathematics* (pp. 1–39). Cham: Springer International Publishing.

Faerber, S. J., Leder, H., Gerger, G., & Carbon, C.-C. (2010). Priming semantic concepts affects the dynamics of aesthetic appreciation. *Acta Psychologica*, *135*(2), 191-200.

Hitsuwari, J., & Nomura, M. (2021). How individual states and traits predict aesthetic appreciation of haiku poetry. *Empirical Studies of the Arts*, *40*(1), 81-99.

Hugentobler, K. G., & Lüdtke, J. (2021). Micropoetry meets neurocognitive poetics: Influence of associations on the reception of poetry. *Frontiers in Psychology*, *12*, 737756.

Jacobs, A. M. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, 186.

Johnson-Laird, P. N., & Oatley, K. (2022). How poetry evokes emotions. *Acta Psychologica*, *224*, 103506.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(Pt 4), 489–508.

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments : The aesthetic episode – developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443-464.

Liu, M. (2022). Mental imagery and poetry. *The Journal of Aesthetics and Art Criticism*, *81*(1), 24-34.

Maruyama, S., & Ishizu, T. (2025). The effects of titles on the aesthetic evaluation of japanese poetry. *Empirical Studies of the Arts*, *43*(1), 707–720.

Mehl, K., Gugliano, M., & Belfi, A. M. (2023). The role of imagery and emotion in the aesthetic appeal of music, poetry, and paintings. *Psychology of Aesthetics, Creativity, and the Arts*, Advance online publication.

Obermeier, C., Kotz, S. A., Jessen, S., Raettig, T., von Koppenfels, M., & Menninghaus, W. (2016). Aesthetic appreciation of poetry correlates with ease of processing in event-related potentials. *Cognitive, Affective & Behavioral Neuroscience*, *16*(2), 362–373.

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., & Kotz, S. A. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, *4*.

Pilkington, A. (2014). Poetic effects: A relevance theory perspective. Routledge.

Rotaru, A. S., Vigliocco, G., & Frank, S. L. (2018). Modeling the structure and dynamics of semantic processing. *Cognitive Science*, *42*, 2890-2917.

Siew, C. S. Q. (2019). spreadr: An r package to simulate spreading activation in a network. *Behavior Research Methods*, *51*(2), 910-929.

Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. *The psychology of humor: Theoretical perspectives and empirical issues*, *1*, 81–100.

Utsumi, A. (2002). Toward a cognitive model of poetic effects in figurative language. *In Proc. of 2002 IEEE International Conference on Systems, Man and Cybernetics*.

Utsumi, A. (2005). The role of feature emergence in metaphor appreciation. *Metaphor and Symbol*, *20*(3), 152-172.

Utsumi, A. (2006). A cognitive approach to poetic effects of rhetorical figures: Toward a unified theory of cognitive rhetoric. *Proceedings of the 19th Congress of the International Association of Empirical Aesthetics (IAEA2006)*, 413-417.

Vitevitch, M. S., Ercal, G., & Adagarla, B. (2011). Simulating retrieval from a highly clustered network: Implications for spoken word recognition. *Frontiers in Psychology*, *2*.