The influence of sensory modality on aesthetic judgments of poetry

Jessica Frame¹, Kendra Mehl¹, Karen Head², and Amy M. Belfi¹

¹Department of Psychological Science, Missouri University of Science and Technology
²Department of English and Technical Communication, Missouri University of Science and Technology

Author Note

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number R15AG075609.

All data presented in the current manuscript can be found at the following OSF repository: https://osf.io/3h5er/

Correspondence concerning this article should be addressed to Amy M. Belfi, Department of Psychological Science, 135 H-SS, 500 W. 14th St., Rolla, MO 65409, amybelfi@mst.edu
Abstract

Poetry can be experienced in multiple sensory modalities – for example, someone might read a written poem (i.e., visual modality) or listen to a spoken poem (i.e., an auditory modality). Readers may also follow along with a written poem while listening to the spoken version, and therefore experience poems in a multimodal manner. Here, we examined whether aesthetic judgments of poems differ based on the sensory modality in which they are experienced. In the present study, participants (N=233) rated three subjective characteristics of poems (vividness of evoked imagery, emotional valence, and emotional arousal), as well as the overall aesthetic appeal of the poems. Participants were randomly assigned to one of the three modalities: text-only (N=81), audio-only (N=74), or combined audio/text (N=78). Our results showed that participants found the audio-only modality to be the least aesthetically appealing, as compared to visual-only and combined audiovisual. Additionally, we found that vividness of imagery was the most important predictor of aesthetic appeal overall (across all three modalities) but also identified a significant interaction between stimulus modality and imagery such that vividness was most important for text-only poems. Finally, we replicated prior work on individual differences in aesthetic appeal, as our results indicate low interrater agreement for aesthetic appeal ratings of poems. These findings contribute to our understanding of the complexity of the aesthetic experience in poetry and highlight the significance of individual differences and the role of modality in its appreciation.

Keywords: poetry; aesthetic judgements; modality; imagery; appreciation
Introduction

Poetry is a complex form of written art that often employs vivid imagery to evoke a wide range of emotions in its readers. The aesthetic experience of poetry is often characterized by its ability to "paint a picture" with words, allowing readers to conjure vivid images in their minds. Or as Horace explained, “As is painting, so is poetry: one piece will strike you more with close inspection, and another if you stand at a greater distance.” For example, reading a poem about a waterfall might cause a reader to really “see” and “hear” it in their mind’s eye, which might then lead them to find the poem highly enjoyable. While historically, the study of poetry has remained solidly in the realm of the humanities, more recently, psychologists have begun looking at poetry as an object of scientific study. Generally, investigations of poetry from the psychology of aesthetics have taken two approaches: On the one hand, researchers may take a stimulus-driven approach to investigating the components of a poem that contribute to its aesthetic appeal. For example, previous work has investigated many of these factors, which often include features such as lexical or sublexical content, rhyme, and meter (Aryani et al., 2016, 2016; Johnson-Laird & Oatley, 2022; Obermeier et al., 2013, 2015, 2015). Such approaches have been considered to be stimulus-driven, such that they focus on the contribution of objective, stimulus-level features to a reader’s aesthetic response.

A second, yet complementary, approach has been to study the subjective features that influence a poem’s aesthetic appeal. That is, while the stimulus-driven, objective features of a poem consist of things like acoustical or lexical features, subjective features are properties of the poem as subjectively perceived by the observer. Our previous work investigated how these subjective features affect the aesthetic appeal in pieces of poetry (Masked Citation). When participants rated poems (haiku and sonnets) on the vividness of visual imagery, emotional
valence, and emotional arousal, our previous research found that vividness of visual imagery was the most important predictor of aesthetic appeal, more so than emotional valence and arousal. Other relevant work has found that subjective perceptions of a poem’s style, such as whether the poem is perceived to have rhythm and can be read fluently, are important predictors of its aesthetic appeal (Ludtke et al., 2014). More recently, we identified that vividness of visual imagery was an even more important predictor for poetry than for other aesthetic stimuli, that is, visual arts and music (Masked citations).

These recent findings illustrate that the subjective features that contribute to aesthetic appeal can vary across different types of art forms. It is plausible, then, that the features driving aesthetic appeal of an artform may be influenced by the sensory modality in which it is presented. In the case of poetry, there are two primary modes of communication: Poetry can either be read (i.e., the observer reads the written poem) or heard (i.e., the observer listens to the poem being read out loud), or, in some cases, technologically-generated poems might even require audiences to engage in interactive experiences. In certain situations, such as at a live poetry reading, the audience may do both – that is, listen to the poem being read while following along with the written version. Just as certain features contribute more to the aesthetic appeal of poetry than music, then, certain features may contribute more to the aesthetic appeal of written versus spoken poems. One can imagine that while reading a written poem, the reader may put their own emotional prosody onto the words; while listening to a spoken poem, the orator imparts their own interpretation onto the audience. The main goal of the present study was to explore this question. That is, we sought to investigate whether there are differences in the subjective features that contribute to the aesthetic appeal of poetry based on its sensory modality.
Most prior work investigating aesthetic responses to poems has focused on written poetry (Aryani et al., 2016; Belfi et al., 2018; Hitsuwari & Nomura, 2022; Kraxenberger & Menninghaus, 2017; Menninghaus & Wallot, 2021). However, there has been some previous research that looked at spoken poetry. In one study, participants listened to either original poems or poems which had been modified to be low in parallelistic features, and rated their emotional responses to and the aesthetic appeal of the poems: Participants liked the original poems more than the modified versions, indicating that parallelistic diction is an important component driving the aesthetic appeal of poetry (Menninghaus et al., 2017). Another recent study used spoken poetry to look at the impact of continuous self-reporting on aesthetic evaluation and emotional responses (Wagner et al., 2021). They found that continuous self-reporting did not seem to affect how an individual rates spoken poems overall, as retrospective ratings of the poems did not differ between individuals who did versus did not make continuous ratings. Their work, however, did not explore the possibility of differences between spoken and written forms of poetry. Our current work aims to look at, when self-reporting their aesthetic evaluation and emotional responses from a poetic piece, whether or not there are significant differences between ratings when changing the modality of how the poetic piece is delivered.

Another recent study explored the relationship between melody and aesthetic evaluation in spoken poems and songs (Scharinger et al., 2023). This study took a stimulus-driven approach by investigating the role of acoustical features in the aesthetic appeal of spoken poems and songs. In this study, audience members attended a concert in which a series of poems were both recited orally and sung with a piano accompaniment. The authors analyzed acoustical features of the spoken and sung poems to obtain various measures in order to identify relationships between perceived melodiousness and acoustic features. They found that subjective ratings of
melodiousness did correlate with several of the acoustic properties of both spoken and sung poems, and suggest that melodiousness is one important component of the aesthetic enjoyment of both songs and spoken poetry. While their work was integral in understanding the importance of certain factors and their contribution to the aesthetic appeal between music and spoken poetry, they only focused on auditory features. Thus far, there has yet to be any comparisons between written and spoken poetry to see if the modality, in general, has any significant effect on an individual’s aesthetic judgement of the piece. Furthermore, it is unknown whether combining audio and visual presentations of poetry would have any influence beyond each sensory modality in isolation.

While there has not been any work comparing written to spoken poems, or the effect of combining these two modalities, research in music has provided some hints about the differences between audio only and combined audiovisual presentations of musical performances. Prior work has looked at the effect of audio-visual information on the aesthetic and emotional appeal of musical performances (e.g., Chiba et al., 2023; Coutinho & Scherer, 2017; Pan et al., 2019; Tsay, 2013). A number of these findings suggest that the combination of audio and visual information can enhance emotional perception when compared to audio-only or visual-only presentations (Thompson et al., 2008). Musicians and non-musicians alike are both shown to rely heavily on visual information when judging aesthetic appeal from music performances (Tsay, 2013), although this original finding has been given more nuanced interpretation given more recent results.

That is, more recent data has indicated that when looking at musical performers of a similar quality, visual cues have been shown to help enhance and differentiate aesthetic judgements when in a competitive context (Chiba et al., 2023). When looking at visual
information from musical performances, performances with high visual expressiveness were rated to be more engaging, enjoyable, and emotionally expressive when compared to performances with low visual expressiveness (Thompson et al., 2008). This research in music has demonstrated that integration of audio and visual information can enhance the emotional impact of music when compared to audio-only musical stimuli (Pan et al., 2019). This research therefore suggests that the combination of auditory and visual stimuli can influence one’s interpretation of a musical performance. While there has been quite a bit of work on this question in music, the current study sought to investigate the potential influence of sensory modality in poetry. And although this is close to an analogy to what we aim to look at in the current study, it is important to note that the true analogous situation would be where participants are listening to the music while reading along with the musical notation, of which there has not been prior work.

In the centuries before print, poets relied on the oral tradition to help disseminate their work. Poetry (especially fixed forms like a sonnet), while complex, are meant to be heard, and the inherent structure of form helps the listener make meaning and remember lines. Even when access to print was no longer an impediment, the emphasis on hearing poetry as a way to best understand it continued to be a critical element of the experience. T.S. Eliot described poetry as “one person talking to another.” Arguably, 21st century audiences are more attentive to the aurality of poetry than any group for several generations. The proliferation of performance poetry, coupled with the ubiquity of social media, has created rock-star-like followings for many contemporary spoken-word poets. Often these fans never read the work by their favorite poets (some do not even publish text versions of their poetry) because these fans are accustomed to only listening. Additionally, national programs, like the Poetry Out Loud competition, have encouraged many young readers to memorize and recite poetry—creating a new sound-focused
community. Consequently, many more people now draw a large part of their meaning-making from sound.

Here, we explored whether the stimulus modality of a poem (written, spoken, or combined) influences the aesthetic appeal of that poem. Participants were randomly assigned to one of three groups: listening only, reading only, or a combined condition in which they simultaneously read and listened to poems. Our first goal was to determine whether the modality in which a poem is presented influences its aesthetic appeal. The second goal of this study was to evaluate the relationships between the vividness of imagery, emotional valence, emotional arousal, and the aesthetic appeal of a poem, and whether this varies based on stimulus modality. Finally, we sought to investigate whether agreement among participants differed between these three conditions (written, spoken, or combined). Our prior work found that interrater agreement was quite low when rating the aesthetic appeal of written poems (Masked Citation), but it may be because with written poetry, each reader puts their own interpretation onto the text. Therefore, we predicted that spoken poetry may show more agreement among raters than written.

Methods

Participants

Participants were recruited in two ways: Online using Prolific (N=140), an online platform for conducting psychological studies (Palan & Schitter, 2018) and online through Sona Systems for the undergraduate psychology pool (N=106). For Prolific participants, we restricted participation to workers who are native English speakers, live in the United States, and have completed at least 100 previous Prolific tasks and obtained approval ratings of at least 95%. A total of 246 participants completed the task. Eight participants were removed for correctly identifying one of the poems by name (at the end of the task, participants were asked “Did you
recognize any of the poems? If so, which one(s)?” and typed their response into a text box. Five participants were removed for being non-native English speakers. This left a total of 233 participants (135 M, 94 F, 4 NB). Participants were an average of 30.22 years old (SD=13.53). Participants were randomly assigned to one of three groups: audio only (N=74; 50 M, 23 F, 1 NB), text only (N=81; 42 M, 38 F, 1 NB) or combined audio/text (N=78; 43 M, 33 F, 2 NB). A one-way ANOVA revealed that there were no differences between the three groups in terms of age, F(2,229)=0.12, p=0.88.

**Stimuli**

The poems selected for this project consisted of 16 sonnets. The sonnet form was selected because of length—14 lines—long enough for participants to get a strong sense of the work, but short enough to keep participants focused. Poems were selected to represent a cross section of contemporary writers and because they would be likely unknown to participants. For the spoken and combined conditions, all the poems were read by the same female reader, a professional poet (co-author KH) in order to create continuity in voice and tone. The reader read the poems in a typical manner of a poetry reading. Recordings were made in a closed sound studio to ensure that each recording was high quality and without any background noise. For a full list of the poems please see Appendix.

**Procedure**

All procedures were conducted in compliance with the American Psychological Association Ethical Principles and were approved by the Institutional Review Board. First, participants read a general description of the study and provided their consent to participate. Participants first read detailed instructions about rating the poems based on modality condition (audio only, text only, or audio and text). After reading the instructions, participants were
presented with each of the 16 poems, in a randomized order, in their given condition (audio only, text only, or audio and text) and rated the stimulus on four dimensions: vividness (“How vivid is the imagery evoked by this poem?”), arousal (“How relaxing or stimulating is poem?”), valence (“How positive or negative is the emotion of this poem?”), and aesthetic appeal (“How aesthetically appealing did you find this poem?”). Each rating was made on a continuous scale using a slider bar; anchors were displayed on the ends of the slider bar. The slider bar started at the middle position of the scale. For vividness, the anchors were “not at all vivid” to “very vivid”; for arousal, the anchors were “very relaxing” to “very stimulating”; for valence, the anchors were “very negative” to “very positive”; and for aesthetic appeal, the anchors were “not at all appealing” to “very appealing.” The range of possible values for each rating was between 0 and 100 in steps of 1, although participants could not see the numerical values of their rating.

After responding to the questions, participants advanced to the next poem; that is, each page only had a single poem. Once finished, participants completed a short demographic questionnaire that asked their gender, age, level of highest education received, and native language. Participants were then asked if they had recognized any of the poems used within the study.

**Analysis**

The main goal of the present study was to investigate which characteristics (vividness, valence, and arousal) best predict aesthetic appeal in poetry and whether there are interactions between these characteristics and the poem modality. To this end, we computed a linear mixed effects model using the lmer function from the lmerTest package (Kuznetsova et al., 2017) in R (version 3.6.2). All ratings were first “centered” around 0, such that the rating scale was transformed from 0 – 100 to -50 – 50. Predictors were then centered within each subject (i.e., “group mean centering”) before being submitted to the model (Enders & Tofghi, 2007). A linear
mixed-effects model analysis was run with aesthetic appeal ratings as the outcome variable, using vividness, arousal, and valence ratings, as well as the interaction of each of these variables with stimulus modality (written, spoken, or combined), as fixed effects. Random effects included random intercepts for subjects and poems, as well as by-subject and by-poem random slopes for the effects of valence, arousal, and vividness. Since stimulus type is a categorical predictor with three levels, contrasts were set using the contr.sum function in R. Pairwise comparisons were calculated using the emmeans (for the main effect of modality) and emtrends (for the interactions between modality and the continuous predictors) functions from the emmeans package in R (Lenth, 2020). Plots were created using the effects package in R.

A secondary goal was to investigate agreement among the participants and whether this differed as a function of the stimulus modality. To this end, we calculated the intraclass correlation coefficient (ICC) from single measures for each rating scale (vividness, valence, arousal, aesthetic appeal) for each modality (written, spoken, combined). To compare ICC values across the three stimulus modalities, we calculated 95% confidence intervals and assessed the degree of overlap between the three modalities (Stolarova et al., 2014). All data presented here can be found at the following OSF repository: https://osf.io/3h5er/

Results

Predictors of aesthetic appeal of audio, text, and combined poetry

First, our model revealed significant main effects of vividness, valence, and arousal (Table 1). Across all three sensory modalities, more vivid imagery, positive valence, and high arousal were associated with greater aesthetic appeal. We also found a main effect of stimulus modality: Overall, audio-only poems were rated as less aesthetically appealing than text-only and combined poems, but there was no difference in the aesthetic appeal of text and combined poems
(Figure 1A). We also identified a significant interaction between stimulus modality and vividness. That is, vividness of evoked imagery was a more important predictor of the aesthetic appeal of text-only poems than audio-only or combined poems (Figure 1B).

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>$\beta$</th>
<th>SE</th>
<th>$t$</th>
<th>$p$</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vividness</td>
<td>0.56</td>
<td>0.02</td>
<td>20.79</td>
<td>&lt;0.001</td>
<td>***</td>
</tr>
<tr>
<td>Valence</td>
<td>0.32</td>
<td>0.02</td>
<td>12.54</td>
<td>&lt;0.001</td>
<td>***</td>
</tr>
<tr>
<td>Arousal</td>
<td>0.06</td>
<td>0.02</td>
<td>2.67</td>
<td>0.01</td>
<td>**</td>
</tr>
</tbody>
</table>

**Main effect of modality:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined vs. Audio</td>
<td>4.32</td>
<td>2.02</td>
<td>2.13</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>Text vs. Audio</td>
<td>5.39</td>
<td>2.00</td>
<td>2.69</td>
<td>0.01</td>
<td>**</td>
</tr>
<tr>
<td>Text vs. Combined</td>
<td>1.07</td>
<td>1.97</td>
<td>0.54</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

**Modality * vividness interaction:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined vs. Audio</td>
<td>0.04</td>
<td>0.05</td>
<td>0.69</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Text vs. Audio</td>
<td>0.14</td>
<td>0.05</td>
<td>2.59</td>
<td>0.02</td>
<td>*</td>
</tr>
<tr>
<td>Text vs. Combined</td>
<td>0.10</td>
<td>0.05</td>
<td>1.90</td>
<td>0.13</td>
<td></td>
</tr>
</tbody>
</table>

**Modality * valence interaction:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined vs. Audio</td>
<td>-0.08</td>
<td>0.05</td>
<td>-1.72</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Text vs. Audio</td>
<td>-0.02</td>
<td>0.05</td>
<td>-0.44</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Text vs. Combined</td>
<td>0.06</td>
<td>0.04</td>
<td>1.34</td>
<td>0.37</td>
<td></td>
</tr>
</tbody>
</table>

**Modality * arousal interaction:**

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined vs. Audio</td>
<td>0.01</td>
<td>0.05</td>
<td>0.34</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Text vs. Audio</td>
<td>0.01</td>
<td>0.05</td>
<td>0.30</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>Text vs. Combined</td>
<td>-0.002</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1.** Results of linear mixed effects model. *$p$<0.05 **$p$<0.01, ***$p$<0.001
**Figure 1.** A) Violin plots depicting aesthetic appeal ratings for audio, combined, and text conditions. Bar indicates mean and standard error (height of bar). B-D) Raw data with overlaid model fits and 95% confidence intervals for vividness (B), valence (C), and arousal (D) as predictors of aesthetic appeal.

**Agreement among raters**

As assessed by ICC (2,1), participants showed similar patterns of agreement across all artistic modalities. Across all modalities, participants showed low-moderate agreement for valence (ICCs ranging from 0.16 to 0.24), and low agreement for vividness (ICCs ranging from
0.05 to 0.12), arousal (ICCs ranging from 0.07 to 0.09), and aesthetic appeal (ICCs ranging from 0.05 to 0.07). In general, the audio group showed lower interrater reliability than the combined and text groups, and across all groups, aesthetic appeal showed the lowest agreement compared to the other rating scales (Figure 2).

**Figure 2.** Intraclass correlation coefficients. Error bars depict 95% confidence intervals.

**Discussion**

In the present work, we aimed to investigate whether the stimulus modality (written, spoken, or combined) of a poem influences an individual’s aesthetic judgment of that poem – that is, do people find a poem more or less aesthetically pleasing based on the sensory modality in which it is presented? Our results indicated that participants found the audio-only modality to
be the least aesthetically appealing, as compared to visual only and combined audiovisual. Additionally, we sought to evaluate the relationships between the vividness of imagery, emotional valence, emotional arousal, and the aesthetic appeal of a poem, and whether these differed for the different sensory modalities. When looking at these relationships, we found that vividness of imagery was the most important predictor of aesthetic appeal overall (which replicates our prior work which used written poems only, Masked Citation). We also identified a significant interaction between stimulus modality and imagery, such that vividness of evoked imagery was a more important predictor of the aesthetic appeal of text-only poems than audio-only poems.

One possible interpretation for why participants rated the audio-only poems as less aesthetically appealing than text-only poems is that spoken poems may not rely on mental imagery as much as written forms of poetry do. When reading a written poem, readers may be more likely to create pictures in their mind; that is, written poetry may require more generative reconstruction on the reader. In contrast, listening to a spoken poem may place the orator’s interpretation on to the listener, which would then allow them to undertake less effort to generate imagery. It is this effort that may therefore influence the aesthetic appeal of written vs. spoken poems. Alternatively, the preference for written poems may have to do with the influence of complexity – that is, it may be the case that more complex poems are more difficult to understand in a spoken-only modality, and the written poem provides the reader with a better “anchor” to focus on to aid understanding. Interestingly, this may relate to research investigating the impact of program notes (in concerts) or descriptions (next to artworks in museums) on aesthetic judgments. In music, prior work has found that program notes can often reduce one’s aesthetic judgment of a piece of music, although this depends on individual factors such as prior
musical training (Bennett & Ginsborg, 2018; Margulis, 2010). In this case, music listeners may seem to prefer putting their own interpretation on the music themselves. For visual arts, prior work has found that the longer people spend reading art labels was associated with greater appreciation of the art (Brieber et al., 2014) and that titles can influence the aesthetic appeal of visual arts (Millis, 2001).

Considering the finding that imagery was more strongly related to aesthetic appeal for written poems, we also speculate that this is due to the generative process required when reading written poems. The text-only conditions may require readers to more actively engage their own imagery while reading each poem. In contrast, when listening to a spoken poem, there are many other aspects that could contribute to aesthetic appeal – for example, the vocal prosody and tone of voice of the speaker. Relatedly, poems being presented in the text-only modality gives readers more freedom to mentally interpret the imagery in their own way rather than another individual doing it for them—thus providing an entirely unique experience for every individual reader based on the imagery evoked by the poem. Related to this, recent research indicated that individuals with better control of auditory imagery experienced more intense emotions when reading poetry, suggesting that auditory imagery ability is critical for the enjoyment of written poetry (Pițur & Miu, 2022).

When looking at interrater agreement, we found that participants had overall low agreement in terms of aesthetic appeal ratings in all three modalities (audio-only, text-only, and combined). Agreement was lowest for aesthetic appeal, followed by arousal, vividness, and more moderate for emotional valence. These results closely replicate our prior work which looked at agreement among raters for poetry (Masked citation) and music (Masked citation). Relatedly, other work has found similarly low agreement values for aesthetic ratings of music (Juslin et al.,
2023), which suggests that this is a highly replicable finding. We can interpret this in a very simple manner: People have their own unique opinions of what is aesthetically pleasing to them. To put it plainly, people have different tastes. The aesthetic experience of poetry, or any aesthetic object, is inherently subjective. Aesthetic appeal of a piece is deemed by an individual based on a variety of factors like personal tastes, preferences, and their own interpretations of the poetic piece. On the other hand, poems can often offer multiple layers of meaning, thus giving readers a large variety of paths to filter down when interpreting a specific piece. The open-ended nature of poetry allows for one individual’s interpretation of a poem to be completely unique.

Our current results are also related to previous work on audiovisual musical performances and their aesthetic appeal. In music, the addition of visual information has been found to enhance the aesthetic appeal of the performance when compared to audio-only music stimuli (Pan et al., 2019), although it is worth noting that the “visual information” in these studies is videos of a performer rather than musical notation, which would be more comparable to the present work. While we saw no difference in aesthetic appeal between the combined and text-only conditions, we saw lower aesthetic appeal in the audio-only condition. Visual cues may therefore enhance the aesthetic appearance of poetry. With the absence of visual information in the auditory-only poems, participants may have been less engaged than in the other conditions where they were actively reading every word in each poem, invoking the cognitive processes to thus create mental pictures that play an important role in poetry’s appeal. The audio-only condition, without the written text, may have therefore felt incomplete to the reader. An alternative explanation could be related to attention - That is, it’s possible that visual information is easier to attend to; when given auditory-only information, listeners’ minds may be more likely to wander. With a written poem, it is possible to return and re-read portions when ones attention may have wandered,
which is not the case with a spoken poem. This may therefore also contribute to the lower aesthetic appeal for auditory-only presentations.

While our study provides valuable insight into the impact of different modalities in the aesthetic appeal of poetry, there are some limitations that must be considered for future research. First, our study was limited to a specific set of stimuli in terms of the poems chosen to be used here. While we did try to encompass a large variety of topics within the set of poems, it would be beneficial to utilize a more diverse range of poetry such as having more variety of poetry genres and types (as prior research has indicated subtle differences in aesthetic ratings of poetry genres, in this case, haiku and sonnet; Masked Citation). Some forms of poetry can also be abstract and may therefore not induce imagery at all, which could be addressed by including more abstract genres of poetry. Another limitation is that we did not ask participants to describe the types of imagery that were evoked. An interesting question for future research would be to identify whether certain types of imagery are more prominent based on the modality of poem delivery. Lastly, our study was limited to only having a single female reader for each poem. Characteristics of the reader could potentially impact how participants rate each poem on a number of dimensions, including its aesthetic appeal. Prior corpus studies of poetry have often used a variety of speakers, including professionals and non-professionals (Menninghaus et al., 2018). We used a single reader for the sake of consistency across the poems and chose a professional poet to maintain realism and ecological validity. Future research could use a variety of different readers with possible variables including vocal prosody, different genders, or whether the reader is a professional poet, to see whether these differences in reader characteristics effect perceptions of poetry.
In addition to limitations with the stimuli, there are important limitations of the measures. For example, the way valence was measured here was regarding the valence of the poem itself, rather than the valence of emotions evoked by the poem. It is possible that the valence of *felt* emotions in a poem has a stronger relationship with its aesthetic appeal than the valence of emotions *expressed* by the poem. One final consideration relates to individual differences – just as each poem might evoke a different degree of imagery, each individual differs in their capability to induce imagery. In the present study, we did not measure such individual differences in imagery ability. In our prior work, we did measure individual differences (using the Vividness of Visual Imagery Questionnaire, VVIQ) and found a significant but small relationship between individual imagery ability and aesthetic ratings of poems, such that individuals with higher VVIQ scores also rated the poetry as overall more appealing (Masked Citation). Additionally, the present study was conducted in a between-subjects manner for practical reasons (i.e., to keep it in a reasonable time frame for online participants) as well as theoretical ones – namely, we were concerned about potential contrast effects of demand characteristics if participants were exposed to all three modalities in a single experiment. However, future research could consider testing the effects of modality on aesthetic appeal in a within-subjects manner.

It is also important to note that our study has practical implications for poets and poetry scholars – for example, in poetry courses, audio versions of poems are often presented to students, or students are assigned to attend live poetry readings. Our findings suggest that making the poem texts available in classrooms and at readings may increase the appeal of these poems. There are also accessibility issues as well: Individuals with hearing impairments or who do not fluently speak the language of a poem would benefit from both auditory and textual
presentations of a poem. Of course, we are not making the claim that one sensory modality is better than another (i.e., that vision is better than audition), but that, in the case of the present results, presenting poetry in an auditory-only modality is associated with lower aesthetic appeal. From reading to oneself or being read to by a narrator, poetry can presented in a variety of modalities. Another form to consider is where poetry is presented as a short film where the narration of the poem is supported by visual animation to aid in conveying a certain theme or idea of the poem. Looking into such demonstrations in future studies would be beneficial to see how these portrayals may perceived in comparison to the other modalities.

In conclusion, this study aimed to understand the influence of different modalities (audio-only, text-only, and combined) on the aesthetic appeal of poetry. In the present work, participants found the audio-only modality to be the least aesthetically appealing, suggesting that including the written text is an important component of a poem’s appeal. This study also reinforced the importance of vividness of imagery as a predictor of aesthetic appeal of a poem, especially in a text-only modality. Furthermore, findings of low interrater agreement in aesthetic appeal replicate prior work both in poetry and music, further highlighting that all sensory modalities are affected by the unique interpretations and tastes that vary between participants. Overall, this study contributes to our understanding the complexity of the aesthetic experience in poetry and highlights the significance of individual differences and the role of modality in its appreciation.
References


https://doi.org/10.1016/j.poetic.2020.101497
Appendix

Table 1 – List of Poems

<table>
<thead>
<tr>
<th>Number</th>
<th>Genre</th>
<th>Poem Title</th>
<th>Poem Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Sonnet</td>
<td>American Sonnet for My Past and Future Assassin</td>
<td>Terrance Hayes</td>
</tr>
<tr>
<td>2</td>
<td>Sonnet</td>
<td>One Hundred Love Sonnets: XVII</td>
<td>Pablo Neruda</td>
</tr>
<tr>
<td>3</td>
<td>Sonnet</td>
<td>Love in a Time of Covid-19</td>
<td>Craig Santos Perez</td>
</tr>
<tr>
<td>4</td>
<td>Sonnet</td>
<td>Instructions on Not Giving Up</td>
<td>Ada Limón</td>
</tr>
<tr>
<td>5</td>
<td>Sonnet</td>
<td>Late Afternoon Stroll on the Cliffs</td>
<td>Laure-Anne Bosselaar</td>
</tr>
<tr>
<td>6</td>
<td>Sonnet</td>
<td>Incantation of the First Order</td>
<td>Rita Dove</td>
</tr>
<tr>
<td>7</td>
<td>Sonnet</td>
<td>Poem Not to Be Read at Your Wedding</td>
<td>Beth Ann Fennelly</td>
</tr>
<tr>
<td>8</td>
<td>Sonnet</td>
<td>On the Cliffs, Newport</td>
<td>Alan Seeger</td>
</tr>
<tr>
<td>9</td>
<td>Sonnet</td>
<td>A Dream Pang</td>
<td>Robert Frost</td>
</tr>
<tr>
<td>10</td>
<td>Sonnet</td>
<td>In Emily Dickinson’s Bedroom</td>
<td>Lloyd Schwartz</td>
</tr>
<tr>
<td>11</td>
<td>Sonnet</td>
<td>Florida Doll Sonnet</td>
<td>Denise Duhamel</td>
</tr>
<tr>
<td>12</td>
<td>Sonnet</td>
<td>Sonnet</td>
<td>Robert Hass</td>
</tr>
<tr>
<td>13</td>
<td>Sonnet</td>
<td>Sonnet</td>
<td>Billy Collins</td>
</tr>
<tr>
<td>14</td>
<td>Sonnet</td>
<td>Unholy Sonnet 13</td>
<td>Mark Jarman</td>
</tr>
<tr>
<td>15</td>
<td>Sonnet</td>
<td>Reckless Sonnet No. 8</td>
<td>Kimiko Hahn</td>
</tr>
<tr>
<td>16</td>
<td>Sonnet</td>
<td>Farm Sonnet</td>
<td>Kitty Carpenter</td>
</tr>
</tbody>
</table>