

Evaluating Second Language Student Poetry: A Study of Instructors

Lara M. Hauer

PhD Candidate

Composition & TESOL

English Department

Indiana University of Pennsylvania

David I. Hanauer

Professor of Applied Linguistics/English

Graduate Program for Composition & TESOL

English Department

Indiana University of Pennsylvania

Lead Assessment Coordinator

SEA-PHAGES Program

Hatfull Laboratory

Pittsburgh Bacteriophage Institute

Abstract

While poetry is used in teaching composition in U.S. university IEP writing courses, methods of evaluating poetry still remain unclear. The aim of this study is to investigate the way college composition instructors evaluate poetry written by second language (L2) student writers in a composition course, and factors influencing their judgment. In a survey, 23 instructors assigned grades to five “non-professional” poems generated from a corpus by Hanauer (2010, 2015a). Subsequently, they provided explicit reasons for assigning those grades, and ranked 10 criteria in order of importance when evaluating the poems. The highest ranked items were imagery, emotional connection, and expressiveness, while the three lowest were grammatical conventions, rhyme, and topic. Similar to previous studies on textual qualities of L2 poetry and beauty judgements, data analysis indicates that instructors tend to approach evaluation of student poetry by considering aesthetic qualities and their own genre understandings. Implications are discussed for instructors incorporating poetry in an ESL or EFL course context.

Keywords: Second language, poetry, evaluation, writing instructors

Increasingly within the framework of the L2 literacy classroom, creative writing and in particular poetry, is being utilized to supplement other literacy tasks (Disney, 2014; Hanauer, 2010, 2011; Iida, 2011, 2012). As with other aspects of the language classroom, students expect to receive feedback and to be evaluated on their poetry writing. However, assessment and poetry writing is often considered to be antithetical; with the former considered judgmental and fixed to established conventions and the latter subjective, personal and offering creative, unique usages of language. If poetry is being used in the language classroom, it is necessary to have some guidelines and appreciation of the way teachers might approach the process of assessing such writing. It is within this context that the current study was developed. The main aim of this study is to understand the way in which professional writing instructors evaluate poetry written by second language student writers and thus to offer some guidelines and understanding of ways in which this form of evaluation could be conducted.

Reading and Evaluating Second Language Poetry

The evaluation of student poetry poses a challenge for instructors of writing and the mechanisms used to evaluate poetry within the language classroom remain relatively underexplored. Bizzaro (1993) has suggested that one of the problems in evaluating student poetry writing might be the lack of consistent standards for the evaluation of poetry. Bizzaro discusses four frameworks that potentially could be applied to evaluating poetry: New Criticism, reader response criticism, deconstruction, and feminism. His preference is for evaluations emerging for New Criticism and reader response criticism. He suggests that response sheet checklists, portfolios and primary trait scoring negotiated between instructors and students in conferences aimed at poetry revision for evaluating creative writing be used for evaluation. Bizzaro states, “I want to provide students with the kind of evaluative material that will enable them to revise their poems later...[and] understand both how they were graded and why they received the grade that they did” (p. 202).

Two articles appearing in *The English Journal* reveal the dilemmas that composition teachers face when evaluating student poetry. LeNoir (2002) acknowledges the discomfort that teachers can face when assigning a grade to poetry that may be deeply personal and the subjectivity that is involved in passing judgment. He suggests maintaining consistency in method, and teacher-student negotiation of rubrics that may include such items as “creativity/originality, imagery, readability/flow, style, detail/development, clarity, mechanical cleanliness, conformity to curricular

requirements (e.g., form), effectiveness/cleverness in use of language and language devices (e.g., simile, metaphor), and complexity of thought” (p. 61). Griswold (2006) also suggests creating assessment lists including “effectiveness (students’ use of specific poetic devices that they are currently learning or have learned), Process (effort, shown through their writing process), and Mechanics” (p. 71). While LeNoir’s and Griswold’s suggestions are useful for evaluating poetry, they were not specifically developed in relation to poetry written in a second language and as such may not relate to this population.

More recently, Iida (2008) has examined poetry assessment in an EFL context and has designed analytic rating scales with a rubric for assessing haiku (p. 178). He calls attention to the “limited number of evaluation systems for haiku poetry” as reported “in academic journals in TESOL and even other fields” (p. 174). Addressing this need, his analytic scale contains categories including personal voice, audience awareness, organization, haiku conventions, and L2 linguistic conventions, overall focusing on a humanistic approach to evaluation.

While not directly a study of assessment, Hanauer (2015) has studied how TESOL and writing graduate students make judgments about the beauty of second language poetry. In this framework, a decision about the beauty of a poem is in itself an evaluative aesthetic judgment or in other words, an assessment. Within this study, the beauty of a poem was empirically related to the reader’s evaluation of how well the poem was crafted and the degree to which this elicited an emotional response. Specifically, increased levels of perception of the quality of the writing and increased levels of emotional response translated into increased rating of the beauty of a particular poem. The whole of the judgment of the aesthetic qualities of the poem was situated within a causal relationship in which the social sanctioning of the writer as a poet enhanced the degree to which the poem was considered to have been well written and able to elicit an emotional response, which in turn increased the beauty of the poem. The results of this study are interesting in terms of assessment in that they both point out the core features readers are responding to (perception of how well the poem is written and its degree of emotional salience) as well as suggesting that if an instructor responded positively to a poem this would change the students’ aesthetic evaluations of their own and others poetry.

Overall, the existing scholarship on ways of evaluating student poetry in the language classroom is limited. Within the creative writing community, as represented in the remarks of Bizarro (1993), there is an assumption of the difficulty of conducting poetry writing assessment.

Applied linguists such as Hanauer and Iida see the option of reaching some guidelines on assessing poetry. Perhaps, the most promising direction is hinted at within Hanauer's (2010) approach to poetry pedagogy in which the instructor responds to student poetry as a reader. In this approach, the features of poetry reading combine with the responses to poetry of an informed reader (as explicated in Hanauer 2015) to evaluate the poem.

Methods

Participants

Twenty-three professional writing instructors completed an online survey (Qualtrics). Instructors were contacted through an email. The majority of participants (78%) were first language (L1) English speakers. The remainder consisted of native speakers of Chinese, Ghanaian, Thai, and Japanese. All participants had graduate level education in North America; the ages ranged from 18 to 50; and 65% identified as female (with 45% identifying as male). On a simple yes/no question, all participants indicated that they considered themselves competent to rate non-professional, second language poetry. All data were collected from participants in accordance with and under the supervision of the Indiana University of Pennsylvania's IRB board (Log #16-250).

Data Collection Instrument and Process

The survey instrument comprised a task grading five L2 student (non-professional) poems written for a college composition course, explaining the reasons for the grade they assigned and ranking 10 criteria in terms of their importance to the rater in evaluation of the five poems. The specific instruction for the ranking task was: "*Think about the poems you just rated. Please rank the following criteria in terms of their importance in making your choice, with 1 being very important and 10 being not important.*" The criteria consisted of the following central components of poetry taught within most educational language programs: imagery, emotional connection, figurative language, form/structure, grammatical conventions, expressiveness, inspiring new thought, word choice, rhyme/sound, and poem topic. The poems used in this study came from Hanauer's (2010) corpus of second language student poetry. These poems had been used in the previous study of beauty judgements (Hanauer, 2015). The poems were randomly chosen from a corpus of 1,000 second language poems (Hanauer, 2015) and have been characterized as short, imagistic poems which elicit emotional responses.

Analytical Approach

The main aim of this study is to understand the way in which professional writing instructors evaluate poetry written by L2 student writers. The central aspect of this analysis is calculating the rankings of importance of 10 features of poetry reading used in judging the quality of a student poem. However, average rank orders will only be significant if there is agreement among raters on the rank orders. Accordingly, as a first stage, an inter-class correlation (ICC) was calculated to assess the degree of agreement among the 23 raters. If high levels of agreement are achieved on an analysis of this type, the conclusion is that the raters are using the different features of poetry reading in a similar way in order to rate the quality of poetry written by L2 writers. Once inter-rater reliability has been established, average rank orders of poetry reading features can then be established. To further validate and understand the instructors' evaluations of poetry written by second language writers, two additional analytical steps were taken. A computational linguistic approach consisting of a simple word counts was used to validate the poetry reading feature rankings. The program TextAnalyzer was used to find high frequency words in the instructors' responses. Finally, instructors' written explanations were analyzed thematically using a content analysis approach to further explicate and understand the ways in which instructors understood and utilized the different features of poetry reading in their evaluations. The thematic analysis was by 2 raters in accordance with guidelines of a content analysis approach.

Results

As explained in the Methods section, the first issue to be dealt with empirically is the degree to which there is inter-rater reliability in the ratings of the 10 features of poetry reading. Since the survey involved rank ordering of 10 poetry reading features by the same 23 participants, a two-way mixed, Interclass Correlation Coefficient with mean ratings and absolute agreement was calculated. A very high degree of reliability was found between instructors on the ranking of the poetry reading features. The average measure ICC was .968 with a 95% confidence interval from .931 to .991 [$F(9,198) = 28.55, p < .0001$]. The outcome of the ICC suggests that the raters were in high agreement over the rank ordering of the poetry reading features that they use in evaluating poetry written by L2 writers.

Since high levels of inter-rater agreement were found, the average rank order of items can be established. Table 1 presents the average rank order for each of the poetry reading features measured in the current study. As can be seen, the top three ranking features consist of imagery,

emotional connections and expressiveness, suggesting that readers are responding to the genre conventions of actually reading a poem rather than as a language-orientated writing task. To further support this interpretation, it is interesting to note that both topic and grammar, which are usually central features of writing evaluation, were considered the least important among all the ranked poetry reading features. Overall, the data presented here suggest that the instructors are responding to the genre of the writing and evaluating features considered important for this type of writing.

Table 1

Mean rank rating for 10 poetry reading features used in the evaluation of poetry written by second language writers

Poetry Reading Feature	Rank Order	Mean Ranking	Standard Deviation
Imagery	1	2	1.41
Emotional Connection	2	2.69	1.76
Expressiveness	3	3.82	1.58
Figurative Language	4	4.56	2.1
Inspiring New Thought	5	5.3	2.45
Word Choice	6	6.04	2.03
Form and Structure	7	6.08	1.92
Rhyme and Sound	8	7.6	2.29
Topic	9	7.95	1.94
Grammar	10	8.91	1.59

In order to further validate the findings of the rank order analysis, the written explanations of reasons instructors gave for the evaluations of the poems they read were analyzed in terms of frequency of word counts. Table 2 presents the highest frequency content word counts. The basic assumption of a validation of this kind is that the rankings made by the participants should be reflected in the open-ended written explanations of how they evaluated the poems they read. As can be seen in Table 2, this is indeed the case. The highest frequency words appearing 13 times or more

reflect issues of positive feeling and imagery. Importantly, within this list, instructors explained their evaluations in terms of connection, beauty, emotion and creativity. All these words validate the rank order data and also suggest that the instructors were responding to the genre of poetry as actual poetry readers.

Table 2

Highly frequency words appearing in instructors' written explanations of the reasons for providing a specific evaluation of a poem written by a second language writer

Word	Frequency
<i>Like</i>	29
<i>Feeling</i>	17
<i>Nice</i>	15
<i>Imagery</i>	15
<i>Image</i>	14
<i>Good</i>	14
<i>Feel</i>	13
<i>Clear</i>	9
<i>Sounds</i>	9
<i>Structure</i>	9
<i>Hope</i>	8
<i>Heart</i>	8
<i>Tense</i>	7
<i>Liked</i>	7
<i>Images</i>	7
<i>Connection</i>	6
<i>Beautiful</i>	6
<i>Emotion</i>	6
<i>Creative</i>	6

The final analysis that needs to be presented deals with the qualitative understanding of the different ranks presented. In order to explore the instructors' understandings, we conducted a thematic analysis of the written explanation of how the poetry evaluations. The results of this analysis are in Table 3. As can be seen in the instructors' responses, the different features of poetry are used for both positive and negative evaluations of the poetry that they read.

Table 3

Instructor statements explaining poetry evaluations organized in thematic categories

Thematic Category	Instructor Statements
<i>Imagery</i>	<ul style="list-style-type: none"> • The primary reason why I assigned A grade for this poem is that the poet could successfully describe the image of piano and her feeling. • I like the poem's central image of an empty frame. • The imagery is very clear with this poem. • Imagery is not especially original. • Simple imagery- has potential, but too trite as is. • This poem's imagery was nice and seemed to be a tribute to a lost love or family member. • The imagery of sights and sounds is effective. • I'd encourage this student to use more descriptive terms, to really invoke images. Ex: "The smell of beer"... what's that really smell like? • Only an A? This is great...tons of imagery and meaning to that imagery

<p><i>Emotional Connection</i></p>	<ul style="list-style-type: none"> • Poem 1 was intriguing. I think it had a very interesting way of guiding readers through the speakers' emotions. • I can see some emotions and descriptions here, but it needs a bit more to engage the readers. • Lots of good emotion here and description. • This is a poem with emotion and as a reader, I can feel that. Also, I can see the student poet's motions in the first stanza like I was there with him/her. • This poem feels unfinished - we have the image of a piano that makes music but there is no energy here. No emotion. • I enjoyed the description of this poem. I could also understand and feel the feeling of emptiness in this poem. • Clear but I suggest add more feeling • Evocative. The two stanzas evoke competing/contrasting images between the middle line. There is nice symmetry to this poem. • Powerful sentiment. • Wow - this poem is very personal and powerful. The feeling of hope shines through the description of the poem.
<p><i>Expressiveness</i></p>	<ul style="list-style-type: none"> • I think this poems' speaker displays a confidence that could be interpreted by readers as positive and motivating. • The speaker is honest with the thoughts he shares both about what is happening now and what he remembers from the past. • This does a nice job of creating a setting and implying something about the speaker's sense of self and involvement. The back story is insufficient to support the final emphasis on dancing. • I like how the third stanza uses the future tense and expresses hopes and wishes • This poem helps the readers engage in the speakers' thoughts.
<p><i>Figurative Language</i></p>	<ul style="list-style-type: none"> • Limited use of poetic devices • I also liked that this poet could use metaphor. For example, I interpreted that "snow-white" indicates not only the color of the keyboard, but also her/his feeling when she touches the keyboard (maybe, the keyboard was a bit cold). Also, s/h used "light and bright" to describe her/his feeling when make some sounds.
<p><i>Inspiring New Thought</i></p>	<ul style="list-style-type: none"> • The images were sort of cliché and basic in the opening lines. • I gave this poem an "A" because it seems to utilize some rhetorical modes that can encourage thinking amongst the readers. • Surprising and fresh message at the end. • Prior to the "I" entering the poem the descriptive terms feel cliché. • Some fresh ideas and nice focus on a key moment/experience.

<p><i>Word Choice</i></p>	<ul style="list-style-type: none"> • Word choice is good, also like the adverb use in the poem. • I love the use of pronouns. • Why "lingers"? This choice of word seems to contradict "cut from albums" because lingers indicates unwelcome or neutral feelings, but "cut from albums" indicates a willful act to extricate pain. • A bit heavy handed with adverbs and adjectives. • What I did not like about this poem is s/he uses "you." I personally interpreted that the loss of somebody important is associated with the poet, and not with me as a reader.
<p><i>Form and Structure</i></p>	<ul style="list-style-type: none"> • The way the last few lines shift across the page provides readers with the illusion of the speaker "letting go" of who he/she is describing. • I like what the student has done with the use of space/white space on the page and enjambment and how that form complements the content. • I like that it is succinct and cohesive. I appreciate the writer's play with the line structure but I don't know that I derive any particular meaning from it. • I like the structure of this poem. For instance, the usage of sentence initial lower case (e.g., your image, or, lingers, or in my heart, etc.) indicates that the sentence is still ongoing even when that sentence is located in a new line. Also, I liked hr/his aesthetic structure (e.g., location of "lingers," "or," :in my heart," "a cut," and "empty frame"). This indicates that loss of picture indicates loss of somebody important for the poet. • I like that the line structure seems to reflect the movement of the person/people dancing. • Effective spacing and line breaks and sparsity of language near the end. • It's not clear to me why this poem is structured the way it is. • Why the weird spacing? Let the words do the work. • Structural issues make the message a bit unclear • I like the way the author staggers the lines to show movement, which reminds me in a way of dancing.
<p><i>Rhyme and Sound</i></p>	<ul style="list-style-type: none"> • I would love to "hear" the sound of the piano that this author feels light and bright. • Nice sounds. • Almost too stilted...whether the meter or the rhythm...sounds very choppy.
<p><i>Topic</i></p>	<ul style="list-style-type: none"> • It's a tough thing to write about but the second half is touching because it talks about the desire to get things back to normal • Why title it the dance when the poem is about movement of dancing not the idea of dancing. The article "the" points to the wrong thing. • I like the topic - it's ripe for expression.

<i>Grammar</i>	<ul style="list-style-type: none"> • Grammar is awkward. • The image of the adult holding the child's hand needs to be revised; it's not the image the writer wishes to describe because of the grammatical ambiguity. • The tense shifts from first person to third person to second person in each stanza is very nice. • The grade for this poem is a low C, maybe even D. The verb tense shift is confusing. Tense shift can be done but the tense is not consistent with what is happening in the poem. • Passive voice doesn't feel like right choice in 2nd stanza • For me, the switch in tenses is very eloquent because it's a flashback technique, almost a PTSD-type flashback style of conveying emotion.... • Why the change in tense?
----------------	---

Discussion

The central aim of this study was to understand the ways in which professional writing instructors evaluate poetry written by L2 student writers. The results of this study offer an answer to this question. The instructors in this study were consistent in their agreement over the ranking of the features important for assessing poetry. Importantly, the three highest ranking features of poetry reading used in evaluating second language poetry consisted of imagery, emotional connection and expressiveness. These features echo work done in previous studies on both the textual qualities of poetry written by second language writers (Hanauer, 2010) and the way in which beauty judgments are made (Hanauer, 2015a). Poetry written by second language writers using the writing approach outlined in Hanauer (2010) are characterized by the presence of imagery, personal voice and emotion (Hanauer, 2010, 2015b). Furthermore, beauty judgments for this type of second language poetry focus on the perceived quality of writing and the emotional content of the poem. Based on the results presented here, it seems that the instructors' way of approaching the evaluation of the student poetry they were presented with was to consider the aesthetic aspects of the poem and to read this poem within the genre conventions of poetry.

In recent years, there have been calls to include poetry writing in the second language classroom (Disney, 2014; Garvin, 2013; Hanauer, 2010, 2011, 2014; Liao, 2016). These calls emphasize the potential of poetry to offer genre diversity as well as ownership and engagement with personal writing in a second language. While these are admirable aims in themselves, most educational settings require student work to be evaluated and graded. The study presented here aims to interact with this issue. It seems that the instructors do have a way of approaching the evaluation

of second language student poetry writing. They seem to address these poems as poems and themselves as poetry readers. This approach prioritizes the features of poetry used to make an aesthetic judgment of a poem; but situates these features within an evaluative framework. Consideration of Table 4 which presents the instructors' explanations of the usage of the different features clearly shows how these elements can be used in an evaluative way for both positive and negative conclusions. It is also important that while none of these teachers had had direct instruction in the assessment of second language poetry writing, there was extensive agreement among the instructors on what was important to consider.

Previous research (Garvin, 2013; Hanauer, 2014, 2015; Iida, 2008, 2011; Liao, 2016) has shown that there are benefits to the usage of creative writing in the language classroom, but for this approach to be practical, students do need to receive feedback. The study presented here suggests that poetry written by second language writers can be evaluated by teachers using their own perceptions of poetry reading. This is an aesthetic response tied to the features of poetry reading which offers feedback from within this position of appreciating the beauty of the written poem. If we position the results found here with Bizarro's (1993) suggestion that the aim of assessment in a poetry class is to direct revision, a space can be constructed for the L2 writing instructor working with poetry. In this context, the writing instructor reads the student poetry as a poetry reader, makes an aesthetic judgment, considers and evaluates the usage of the poetic features and integrates this both within the assignment of a grade and as a point of discussion with the student for revision. Language and beauty, poetry and writing become the center of discussion offering a very different type of student-teacher interaction. As suggested by Iida (2008) and Hanauer (2014), this suggests a very different orientation for the teacher of writing than that which is usually experienced in the L2 classroom.

In some ways, the conclusions of this study change the self-positioning of a writing instructor in the same ways as the writing of poetry changes the positioning of the student writer. Poetry evaluation allows the teacher the option of actually interacting with the emotive and communicative content of the student writers and to offer feedback dealing with the concept of beauty in language. There is a freedom for the teacher in a position of this kind, and no doubt student conferencing around evaluations of this type would be radically different from other interactions present within the language classroom. It is with the hope of facilitating discussion and interactions dealing with aesthetics around poetic writing while maintaining the option of assessment that the current study was conducted.

References

- Bizzaro, P. (1993). Literary theory, composition theory, and the reading of poetry writing. In *Responding to student poems: Applications of critical theory* (pp. 1-13). Urbana, IL: NCTE.
- Disney, D. (Ed.). (2014). *Exploring second language creative writing: Beyond Babel*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Garvin, R. T. (2013). Researching Chinese history and culture through poetry writing in an EFL composition class. *L2 Journal*, 5(1), 76-94.
- Griswold, A. (2006). Assessment lists: One solution for evaluating student poetry. *The English Journal*, 96(1), 70-75.
- Hanauer, D.I. (2010). *Poetry as research: Exploring second language poetry writing*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hanauer, D. I. (2011). The scientific study of poetic writing. *Scientific study of literature*, 1(1), 79-87.
- Hanauer, D. I. (2014). Appreciating the beauty of second language poetry writing. In D. Disney (Ed.), *Exploring second language creative writing: Beyond Babel* (pp. 11-22). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Hanauer, D.I. (2015a). Beauty judgments of non-professional poetry. *Scientific Study of Literature*, 5(2), 283-199.
- Hanauer, D. I. (2015b). Measuring voice in poetry written by second language learners. *Written Communication*, 32(1), 66-86.
- Iida, A. (2008). Poetry writing as expressive pedagogy in an EFL context: Identifying possible assessment tools for haiku poetry in EFL freshman college writing. *Assessing Writing*, 13, 171-179.
- Iida, A. (2011). *Revisiting haiku: The contribution of composing haiku to L2 academic literacy development* (Doctoral dissertation). Retrieved from <http://knowledge.library.iup.edu/etd>.
- LeNoir, W.D. (2002). Grading student poetry: A few words from the devil's advocate. *The English Journal*, 91(3), 59-63.
- Liao, FY. (2016). Identities in an ESL poetry book: A case study of one Chinese student. *The Journal of Literature in Language Teaching*, 5(2), 46-61.

Appendix A: Sample Poem

Piano

The piano shined
 black and heavy
Keyboard
 snow-white,
Covered with a dark red cloth.
I washed my hands cleanly
I touched the keyboard
 nervously.
The sounds ring
 in my small room
My heart sounds,
 light and bright.