The Relationship Between Textual and Reader Variables in Literary

Perception: An Empirical Approach to the Reception of Japanese

Poetry in Translation

James Michael Garza

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

University of Leeds

School of Languages, Cultures and Societies

January 2021

The candidate confirms that the work submitted is his own and that appropriate credit

has been given where reference has been made to the work of others.

ACKNOWLEDGEMENTS

ABSTRACT

The question of what makes a text 'literary' has been philosophically evergreen, stretching back to Aristotle's *Poetics* and weathering every critical approach through the modern era. Since the 1970s, researchers in the field of Empirical Studies in Literature (ESL) have tried to determine whether literary processing is principally text-directed (the formalist position) or reader-directed (the conventionalist position), based on observations of actual readers. Given the widespread characterisation of reading as intensely subjective, we might expect to see a huge diversity of response—yet major studies in ESL have reported results in favour of the formalist position. In expanding its sample beyond undergraduates, this survey-based study of Japanese poetry in English translation gives diversity of response a greater chance to emerge, and produces evidence suggesting that the case for the formalist position is not as strong as previously thought. Moreover, this thesis engages with recent work in Translation Studies to account for the translated-ness of its stimulus texts. Recognising translation as an important litmus test for understanding 'literariness,' this thesis operationalises a major contemporary theory of translation for the first time: Venuti's ([1995] 2008) theory of foreignisation. Against its proponent's objections, I make the case for the amenability of this theory to empirical testing, laying the groundwork for further study. Building on Eco (1990), I advance a useful new concept (*intentio translatoris*) to schematise the dynamics of reader response in Venuti's theory and explain the shifting criteria by which he describes translations as foreignising. Following an in-depth comparison between Russian Formalism and Venuti's foreignisation, I conclude that the latter theory does not totally escape the instrumental model it rejects. I do this to challenge unspoken assumptions about the translator as source of invariance, and to encourage a reassessment of the role of *intentio lectoris* in reading literature in translation.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ABBREVIATIONS AND SYMBOLS

ACC = accusative

ADJP = adjectival phrase

BG = background

BNC = British National Corpus

COP = copula

COMP = complementizer

CONJ = conjunction

D = determiner

DAT = dative

ESL = Empirical Studies of Literature

EXC = exclamatory

FG = foreground, foregrounding

FID = free indirect discourse

GEN = genitive

H = head

H_a = alternative hypothesis

H_0 = null hypothesis

IMP = imperative

INT = interrogative

LOC = locative

NEG = negation

NOM = nominative

NP = noun phrase

P = preposition

PCA = principal components analysis

PP = prepositional phrase

PH = pre-head

S = sentence

SUP = suppositional

T = tail

TP = tense phrase

TOP = topic

TS = Translation Studies

TS = tonic syllable (in intonation analysis)

VP = verb phrase

Note: I use double inverted commas (" ") for quotations from other sources, and single inverted commas ('') when drawing attention to a word or phrase for reasons other than academic citation (such as signalling distance from contested terminology). While I follow British spelling conventions, some of the sources I cite use American spelling. I note this here to avoid using 'sic' in every instance. All quoted italics/emphases are present in the original unless otherwise noted.

1. INTRODUCTION

1.1. GENERAL INTRODUCTION

One of the central puzzles with which this thesis is concerned is the seeming mismatch between humanistic and empirical accounts of literary reading. *If the experience of reading poetry is deeply subjective, then how do we account for the apparent lack of diversity of response reported in some empirical studies?* As a Translation Studies researcher interested in this question, I was forced to reckon with the fact that most studies in the field of ESL (Empirical Studies of Literature) have concerned themselves with 'original' (that is to say, non-translated) texts, while those that *have* employed translated texts have tended to leave unexplored the theoretical challenges that translation poses to the above question. This orientation informs the second main question of this thesis: *How does the foreknowledge of the translated-ness of texts participate in the affective and cognitive experiences of target-language (TL) readers of translated poetry?* Finally, the choice of language pair investigated in this thesis (Japanese/English) reflects my concern as a translator of Japanese poetry with the theory/practice divide—in particular, the question of how translation theorists like Lawrence Venuti (see e.g. Venuti [1995] 2008) justify the move from descriptive research to recommendations for future practice,[1] and how such theories might inform my own translational practices.

To answer these questions, I designed a cross-sectional, survey-based study of real readers' responses to translated Japanese poetry. These surveys were designed to collect data to test predictions about the nature of poetry reading and the effect of translation features described by Venuti as successfully foreignising in Anglophone reception contexts. In

---

[1] This question was posed to me at an early stage by Jean Boase-Beier (email to author, 19 November, 2014).

designing the surveys and interpreting results, I engaged deeply with the following issues:

1) what the Russian Formalists called *literaturnost* (literariness), or "that which makes of a given work a work of literature" (Jakobson 1921: 11; qtd. in Erlich 1980: 172)

2) what the Prague Structuralists called *aktualisace* (foregrounding), or "the violation of the norm of the standard" which "makes possible the poetic utilization of language" (Mukařovský [1932] 1964: 18)

3) what the translation theorist Lawrence Venuti calls "foreignizing" translation, i.e. a translation practice which "seeks to register linguistic and cultural differences" (Venuti [1995] 2018: xiii), and which "deviat[es] enough from native norms to stage an alien reading experience" (Venuti [1995] 2008: 16)

4) how translation and the theoretical perspectives offered by Translation Studies change the way we know, investigate and use categories like 'literary'/'non-literary' and 'standard'/'non-standard'

My thesis offers several original contributions to knowledge. First, it provides evidence which suggests that the case for the text-directedness of foregrounding (FG) is not as strong as previously thought. Second, it develops a methodology for empirically investigating a major contemporary theory of translation (foreignisation), and operationalises this for the first time in the field of ESL. Third, building upon Umberto Eco's work on response, it introduces a useful new concept—*intentio translatoris* [intention of the translator]—to schematise the dynamics of reader response in Venuti's theory of foreignisation and explain the shifting criteria by which Venuti describes translations as foreignising.

In this chapter I discuss the rationale and design of the study, including the stimulus texts and hypotheses tested. I also offer definitions of some key terms. In Chapter 2: Background, I

discuss the historical background of the two competing philosophies of literariness whose claims are at stake, before moving on to a discussion of the previous empirical research. Next I undertake the first substantial comparison of Russian Formalism and Venuti's theory of foreignisation, with a particular emphasis on instrumentalism and constraints on reader response. These are key concepts moving into Chapter 3: Methodology, where I demonstrate the amenability of Venuti's theory to empirical investigation, despite his stated opposition to empirical translation research. As I demonstrate there with my concept of *intentio translatoris*, Venuti's model does not in fact escape the instrumentalism he criticises other translation theories for. Next, I discuss the features Venuti associates with successful foreignisation, as well as my classification of the stimulus texts into two categories ('ostensibly foreignising' and 'ostensibly domesticating') for Survey 2. I close out Chapter 3 with a general description of the stylistic analyses employed for Surveys 1 and 3, and a more detailed description of the samples of respondents. This brings us to Chapter 4: Stylistic Analyses—owing to limitations of space, I can only include one such analysis in full. Nevertheless, I present the relevant data for the other translations afterward. In Chapter 5: Empirical Results and Discussion of Findings, I discuss the results for each hypothesis, as well as the implications of those results for the theories investigated. I expand upon these implications in Chapter 6: General Conclusions, revisiting the questions with which I opened this thesis and positing future research directions.

1.2. PROBLEMS OF RECEPTION

The purpose of my study is to test claims about the nature of poetry reading, with an aim toward clarifying unresolved issues in the field of ESL, as well as advancing the debate about the role of foreignisation in translation (cf. Venuti [1995] 2008). I argue that not only are the

two issues linked, but they present a chance to view one field from the vantage point of the other, to the mutual benefit of both ESL and TS. In collecting data on British readers' responses to translated texts, my research addresses acute problems in both disciplines. These ESL and TS problems hinge on the same variable: the reader. Or to be more precise: *readers* in the plural, as it is precisely the issue of reader diversity (in the sense of diversity of response) that comprises the difficulty for this study. I provide background and context for these problems below, starting with a discussion of the problem of reception in ESL, before moving on to a statement of the problem in TS.

1.2.1. THE PROBLEM OF RECEPTION IN ESL

Among consumers of other cultural products, diversity of response has been long attested by researchers in sociology, media studies, and semiotics, fields with sophisticated frameworks to account for systematic differences in the perception of stimuli like art, television broadcasts and other media. However, when it comes to literary texts, empirical research has yielded inconsistent results about the role of reader background on literary processing. In my view, such research has typically not provided enough opportunity for diversity of response to emerge, owing to methodological issues which I discuss at the end of this section. This is what I mean by the reception problem in ESL.

Influential works that have contributed to the view of audiences as stratified, as opposed to undifferentiated, include Bourdieu's ([1968] 1993) "Outline of a Sociological Theory of Art Perception," Eco's ([1965] 2003) "Towards a Semiotic Inquiry into the Television Message" and Hall's ([1980] 2005) "Encoding/Decoding." All three explain diversity of response in terms of a mismatch between codes: those of the decoder, and those of the encoder. Moreover, in all three explanations, the operation of these codes depends on broader

social conditions: for Bourdieu, acquisition of the necessary codes for art appreciation is determined by "family environment and school" ([1968] 1993: 227ff.). For Eco, decoding depends on the receiver's "general framework of cultural references," which includes "his ideological, ethical, religious standpoints, his psychological attitudes, his tastes, his value systems, etc" (Eco [1965] 2003: 13). Finally, Hall ([1980] 2005: 114-116) argues that viewers' styles of decoding broadcasts depend on their willingness to accept hegemonic accounts of events—that is, to "take the meaning as they—the broadcasters—intended" (ibid.: 114).

Given the myriad similar mismatches we can imagine in encounters between readers and written texts, we might expect to see strong evidence of that outcome so dreaded by the opponents of reader-response criticism: a profusion of individual responses which undermine claims to the definitiveness of expert readings. As Fish (1980a: 4) once summarised this position, the "chief objection to talking about the experience of the reader" has been that "there are (at least potentially) as many experiences as there are readers, and that therefore the decision to focus on the reader is tantamount to giving up the possibility of saying anything that would be of general interest."

Yet empirical research on the effect of reader background on literary perception does not exactly bear this out. Some studies, like Van Peer (1986) and Miall and Kuiken (1994), have reported considerable agreement among readers, and between readers and experimenters, in the perception of literary effects, despite differences in the levels of literary training of respondents. On the other hand, studies like Dixon et al. (1993), Hakemulder (2004) and Zyngier, Van Peer, and Hakemulder (2007) have produced evidence suggesting that the appreciation of more complex texts may depend to some degree on literary socialisation. To understand the import of such findings, we must familiarise ourselves with the goals and theoretical background of such research, and specifically what is meant by the terms 'literary

effects' and 'foregrounding.'

The above researchers and others have sought to empirically study the nature of literary reading. Specifically, they have tested claims relating to two competing schools of thought on literary reading. One view holds that literary processing is principally text-directed, i.e. contingent on the presence of certain textual features. This is the formalist or traditionalist position on literariness. (I use the lower-case 'l' here to distinguish between formalism as a position in ESL and Russian Formalism.) The other view holds that literary reading is principally reader-directed, i.e. contingent on the reader's background and conventions for reading certain text types. This is the conventionalist position.

Thanks to the above empirical studies, among others, we know that responses to literary texts are indeed amenable to empirical testing: we can observe and study the effects of literary texts on real readers, provided we know what to look for. Historically, the theory with the largest influence on researchers' ideas of what to look for has been the theory of foregrounding (which I discuss in detail in Chapter 2). This theory, which the Prague Structuralists developed out of earlier work by the Russian Formalists, and further developed by subsequent theorists in linguistics and stylistics, holds that certain kinds of language use prompt a change in the quality of the reader's attention. While these special kinds of language use can take various forms, for instance the breaking of linguistic rules, or the patterning of sounds in statistically unlikely ways, the crucial point is that they make an appreciable impact on the reader compared to the other language in the text. Of particular interest to experimenters, who after all must postulate variables to use as measures of FG, is this passage from Shklovsky ([1917] 1965: 12): "The technique of art is to make objects 'unfamiliar,' to make forms difficult, to increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged."[2]

---

[2] Shklovsky's term for this effect was *ostranenie*, or "defamiliarization" in Lemon and Reis's (1965) translation.

There are certain ways of measuring or tracking this change in attention quality. One is to have readers themselves indicate it. For instance, in his pioneering study Van Peer (1986) had respondents read four lyric poems in which they were to underline words, phrases or lines they found to be "striking," a word chosen to account for the principal effect of foregrounding as discussed in the theoretical literature, while at the same time minimising the potentially confounding effect of personal reactions toward the description of the property being measured (Van Peer 1986: 29). As Van Peer (ibid.) explains:

> It was decided to employ a rather neutral term such as 'striking' (instead of 'significant,' 'unusual,' 'strange,' 'unfamiliar', 'novel,' etc.), on the basis of the fact that a neutral term is more accessible to a range of informants and at the same time may be less parasitic on personal value systems than terms such as 'strange,' 'weird,' etc."

For his main experiment, participants were enlisted from three student populations, forming a "cline of expertise" (ibid.: 45) based on previous training in literature, so that one group was familiar with the theory of FG, one group had a "general knowledge of literary affairs" (ibid.), and one group had "no academic training whatsoever in the reading of literature" (ibid.). All were students at the University of Lancaster. One must hunt around in the data for the total number of students—but this appears to be 52 (ibid.: 121). All were asked to complete a questionnaire about their attitudes toward poetry before moving on to a series of tasks for each poem, including underlining passages for strikingness and discussion value, and ranking lines based on importance to the poem. After statistical analysis of results, Van Peer found that foregrounded passages "were a strong predictor of reader reaction," a result which held true "for all texts, and for all subject-groups, regardless of familiarity with the theory, prior literary training or attitudes towards poetry" (ibid.: 176).

In a study partially inspired by Van Peer's (1986), Miall and Kuiken (1994) obtained

similar findings using modernist short stories as text stimuli. With the aid of computers, they were also able to test another component of the theory of FG, namely that it increases length of perception. The researchers enlisted a total of 198 participants at the University of Alberta, recruiting from "senior level English courses" (Miall and Kuiken 1994: 398) as well as "Introductory Psychology courses" (ibid.: 402). Respondents were tasked with reading a story twice on a computer screen; on the first reading, the story appeared segment by segment, with the reader in charge of when each new segment would appear. In this way, reading times were measured automatically by the computer. On their second reading, respondents were asked to rate each segment according to one of several criteria determined by the experimenters. These criteria included strikingness and affect, a measure of how "evocative of feeling" (ibid.: 393) a particular segment was. Upon analysis of their data, the researchers found that "the degree to which foregrounding is present in the segments of a story is a predictor of both reading times and readers' judgments of strikingness and affect" (ibid.: 404).

While these are major studies in the field, they are not beyond methodological scrutiny. The biggest potential problem is the nature of their samples. This is less an issue of sample size than make-up: the respondents were all students. In fact, respondent groups have consisted exclusively of undergraduates in all but two of the empirical studies I cite in this paper (those are Zyngier, Van Peer and Hakemulder [2007] and Belfi, Vessel and Starr [2018]). While I am sympathetic to the difficulties of finding non-student participants, the obvious objection to the above samples is that undergraduates may not be the most representative population to generalise from. Moreover, one could argue that the test results reflect reading practices that students have had to internalise *to attend university in the first place*, regardless of their major area of study. Perhaps by the time students reach university, they are already sufficiently trained to respond in a basically similar way.

Thus, one of the most productive questions we can ask in light of Van Peer's (1986) and

Miall and Kuiken's (1994) results is: what would the data look like if we expanded respondent samples beyond undergraduates? It is possible, given Van Peer's expertise and sensitivity to matters both linguistic and statistical, that his results are perfectly valid, but need to include the caveat: *for that population*. When we read Bourdieu (1984), De Certeau (1984), or Fish (1980b), we encounter sophisticated arguments about the influence of social institutions of various kinds in the construction of appropriate attitudes and interpretations. In light of their accounts, it seems a major shortcoming for the field of ESL not to investigate other kinds of readers.

One of the innovations of my study is that it enlists the participation of respondents not normally included in this type of research. I discuss my sampling procedure, and the nature of the sample, in section 3.3. The motivation is to give the diversity of response predicted by the conventionalist position on literariness a fair chance to emerge. For now, it seems that the investigation of the true extent of the influence of reader variables has been constrained by the issues I have discussed above.

1.2.2. THE PROBLEM OF RECEPTION IN TS

The debate over the role of foreignisation in translation has also been prosecuted with scant reference to audience diversity. The issue in question is a pivotal one, to which much political and economic import has been attached: **the problem is the reception of translated works, and the difficulty of formulating translation strategies based on reception outcomes which are highly unpredictable, owing to the diverse make-up and dispositions of audiences (among whom we can count those initial gatekeepers with the material and/or cultural resources to enlarge and shape audiences: literary agents, acquiring editors, publishers, reviewers, etc.).** It is difficult to predict the effect of

translation style, on the success—or lack thereof—of a translation. However, in studying the reactions of readers from a variety of backgrounds, we may test whether features presumed to have certain effects really do. In so doing, this research can provide much needed empirical clarification to vigorously debated claims about translation style.

Generally speaking, this debate concerns the viability of strategies presumed to result in "resistant" (Venuti [1995] 2008: 12), "foreignizing" (ibid.: 19), or "minoritizing" translations (ibid.: 263). While I examine how these terms relate to one another in Chapter 3, they are associated with translation practices which, owing to the choice of text and/or specific translation strategies, are supposed to result in translations that resist "ethnocentrism and racism, cultural narcissism and imperialism" (ibid.: 16) in the receiving culture. As Venuti ([1995] 2008: 125) writes:

> A translator can signal the foreignness of the foreign text, not only by using a discursive strategy that deviates from prevailing discourses (e.g. dense archaism as opposed to transparency dependent on current standard usage), but also by choosing to translate a text that challenges the contemporary canon of foreign literature in the translating language.

At the crux of the debate are two major criticisms: one is the inconsistency in Venuti's applications of the terms 'domesticating' and 'foreignising' (see Baker 2007: 152, Tymoczko 2000: 34-40, and Myskja 2013: 12-13). Venuti has sometimes presented these as merely "heuristic concepts [...] designed to promote thinking and research" (cited in Munday 2008: 145-146), and sometimes as a basis for recommendations for practice. The other major criticism, in its broadest sense, concerns the role readers play in deciding the overall political and social uses of translations, uses which may be in contrast to the translator's stated intentions (see e.g. Shamma 2005).

While I discuss problems of intention (and *intentio translatoris*) in Chapter 3, I will address the above two problems briefly. First, even when we concede Venuti's point that "fluency" and "resistancy" are "culturally variable and historically contingent" (Venuti [1995] 2008: 19), we still have no criteria for determining which of the many concurrent forms of "cultural oppression" a translation should resist, nor "how much resistance" a translation should contain to be deemed foreignising, as Tymoczko (2000: 36-37) argues. Can a translation designed to "deviate from prevailing discourses" (Venuti [1995] 2008: 125) yet which fails to achieve the hoped-for resistant effect *really* be deemed a foreignising translation? Here we begin to see the difficulties faced by translators hoping to achieve the kind of political outcomes described by Venuti. As Myskja (2013: 7) writes, "if one cannot establish what constitutes foreignisation, how can translators then take it in use to achieve the desired resistance?"

Moreover, even small, specialised audiences are likely to disagree on certain matters (take, for instance, professional literary or film critics). This plurality also has a temporal dimension. Myskja (2013: 17) points up the changes that a text's audience can undergo through time: "if the [anti-ethnocentric] effect depends on the readership, the effect can never be settled, since the readership itself is and must be an open category." But even putting the future aside for a moment—for a text to have a foreignising effect in the immediate reception scenario seems a tall enough order. This is especially true when, as Shamma (2005: 63) argues, translations can have effects quite contrary to the translator's "articulated intention," as his analysis of the reception of Burton's *Arabian Nights* shows.

This speaks to the second major criticism of Venuti's theory, that it does not acknowledge the ways in which translations can be put to various political uses, regardless of whether they are domesticating or foreignising, a point both Tymoczko (2000: 35) and Shamma (2005: 65) make. Based on a survey of initial reviews of Burton's *Arabian Nights*,

Shamma contends that although Burton explicitly set out to "preserve the linguistic and cultural difference of the foreign text" (ibid.: 63), his translation generally "restated and perpetuated the Western age-old stereotypes about the East" (ibid.: 61). Even if we are sympathetic to Venuti's rebuttal that the translation *would likely have had* the intended effect on its target audience, who were "primarily an educated elite capable of evaluating his translation" (Venuti [1995] 2008: 272), the question about which kinds of resistance make a text foreignising is still applicable. Why is it the text's foreignising effect among elites, in this case, that makes it a resistant translation, while in other cases, for instance, Megan Backus's translation of Yoshimoto Banana's *Kitchen*, what counts is the text's foreignising effect on a so-called popular readership (Venuti [1995] 2008: 121)?

So far I have mainly discussed the theoretical importance of my research problem. Now I will address its real-world implications. The translator's difficulties are compounded by pressures to conform to "native norms" (Venuti [1995] 2008: 16)—to produce texts that are "eminently readable and therefore consumable on the book market" (ibid.: 12). Venuti contends that in Anglo-American receiving scenarios, translations are "judged acceptable by most publishers, reviewers and readers" when they read "fluently" (ibid.: 1). In support of this claim, he cites numerous reviews in which fluency and related qualities are praised, while the various hallmarks of "translationese" are condemned (ibid.: 2-4). Of the numerous claims Venuti's critics have singled out for criticism, this one seems to have engendered the least opposition. Even Pym (1996: 147), perhaps Venuti's most trenchant critic, is sympathetic on this point: "there is indeed the peculiar convention that our target language should be as neutral as possible. This is close to home: as an Australian I once rendered half a Spanish novel into Australian English (full of 'mates' and 'chooks') but abandoned the project because no one took it seriously."

Thus, the problem I am studying is acute on two levels. The first level is theoretical and

concerns the contemporary state of TS. The second level is practical and concerns the ways literary translators work, potentially affecting their ability to publish and thereby profit (where they can) from their translations.

From the perspective of ESL, it is fortunate that readers should be implicated in this problem. It means the problem is amenable to testing, through the analysis of actual readers' responses to translations embodying different qualities, some more in line with the strategies discussed above, and some less. It is hoped that this attempt to investigate the question from a different angle will advance what is in many ways a stalled debate. I provide an outline of my methodology in the next section.

1.3. OUTLINE OF METHODOLOGY

My project is a survey-based study of readers' responses to translated Japanese poetry. The study is designed to collect data to test predictions about the nature of poetry reading and the effect of translation strategies described by Venuti as successfully foreignising in Anglophone reception scenarios. The survey format has been chosen for several reasons: 1) it can be read at home, or anywhere else respondents choose to complete it, under conditions closer to private or leisure reading, which not only minimises the possible effects of institutional factors on responses but also safeguards against potential fatigue effects; 2) it lends itself to a number of different recruitment methods, including sampling by mail, as well as convenience sampling; and 3) it is the format used in Van Peer's (1986) study, whose general methodology, apart from a few innovations, I have adopted.

The basic method of my survey consists in presenting readers with four texts, and asking them to complete a number of tasks to gauge their responses to these texts. Per Van Peer

(1986: 32), the dependent variable consists of readers' responses to the tasks they are asked to complete, while the independent variable consists of the text stimuli, which appear "under two different conditions: FG or BG […]." The determination of which text elements are foregrounded requires a detailed stylistic analysis of the poems, carried out before the experiment. So far, my experiment is similar to Van Peer's (1986).

However, each text slot in my survey will be occupied by one of two different English translations of the same Japanese poem. Thus, the text corpus consists of eight possible texts. These competing translations have been specifically chosen for the way they conform to descriptions of two different translation styles: domesticating and foreignising (see Venuti [1995] 2008). Each respondent is presented with a survey such that each of the four Japanese poems is represented, albeit in only one of the two possible text conditions, i.e. in either an 'ostensibly domesticating' or 'ostensibly foreignising' translation. As I discuss in Chapter 3, my use of the term 'ostensibly' is a response to Venuti's shifting criteria for designating texts as foreignising or domesticating. Note: in no test situation will respondents be directly comparing translations of the same poem. As I am interested in collecting data about impressions of foreignness, it is imperative to avoid giving respondents access to materials that could lessen the impact of features of the foreignising translations.

In my classifications, I appeal not only to the criteria Venuti generally associates with foreignisation, but also to features singled out in his discussions of successfully foreignising Japanese-to-English translations. These include: **the retaining of Japanese vocabulary** (Venuti [1995] 2008: 121), **the combination of "current standard usage with colloquialisms and poetical archaisms"** (ibid.), **"heterogeneity of theme and style"** (ibid**.), the representation of "social situations that differ markedly from the UK and the US"** (ibid.: 161), **and the inclusion of material that "runs counter to prevalent Orientalist stereotypes"** (ibid.: 121) or **"tampers with the representation of women"**

(ibid.: 162).

The size and nature of the text corpus reflect several concerns. Foremost is the practical difficulty, attested by Van Peer (1986: 59), of carrying out stylistic analyses. As he remarks, it would have been "difficult to work with a large corpus because of the large amount of time involved in preparing full-scale stylistic analyses of even short texts" (ibid.). Next is the necessity to control for fatigue effects. To counterbalance this, Van Peer (1986) chose to work only with texts of "'average' length" (ibid.), and also randomised the order of presentation of the texts (ibid.: 48). The texts I have chosen are typically even shorter than Van Peer's. One reason for using shorter texts is Van Peer's (1986: 60) apprehension that in longer texts foregrounding "might appear in a more diluted form," and therefore its effects "might be more difficult to observe."

It is necessary to carry out a stylistic analysis of the poems to predict which elements might attract extra attention from readers. This analysis takes place on three levels: phonology, syntax and semantics. These are the three levels on which FG has traditionally been theorised to occur (Van Peer 1986: 23). Four major types of FG have also been described in the theoretical literature: internal deviation, determinate deviation, statistical deviation and parallelism. I discuss these in Chapter 3.

Properly speaking, the study consists of three experiments, with the majority of the hypotheses tested in Surveys 1 and 2. The first experiment is designed to overcome some of the methodological limitations discussed earlier, chiefly the issue of respondent groups consisting entirely of students. The second experiment is identical in almost all respects to the first, except that instead of strikingness readers will be asked to underline passages that contribute most to the impression of foreignness. Comparing the responses of these two experiments allows us to determine how (if at all) foreignising effects differ from foregrounding.

The third experiment is similar in design to the first two, but its goal is to investigate whether foreknowledge of the translated-ness of texts affects responses from readers. This can be investigated in a rather straightforward way using the same text stimuli, with only minor changes to the language in the surveys. Readers are presented with one of two versions of the same survey: in one version, the poems are explicitly referred to as translations, while in the other survey, the texts are referred to simply as poems. In all surveys (Surveys 1, 2 and 3), I withheld the names of ST authors and TT translators to avoid introducing familiarity as a confounding variable. Finally, no respondent participated in more than one experiment.

1.4. THE TEXT CORPUS

The text corpus consists of eight English translations. These comprise two translations each of four Japanese poems. I present a table below for ease of reference.

Fig. 1. The Text Corpus

Label	Text	Source Author	ST Pub. Date	Translator	TT Pub. Date	Text Condition
Poem A	Shijimi Clams	Ishigaki Rin	1968	Janine Beichman	2005	Ostensibly Foreignising
Poem B	The Isolation of Two Milliard Light Years	Tanikawa Shuntarō	1952	Bownas and Thwaite	1964	Ostensibly Domesticating
Poem C	No title (*Midaregami* no. 175)	Yosano Akiko	1901	Hiroaki Sato	1981	Ostensibly Foreignising
Poem D	Instant	Yoshihara Sachiko	1964	Miller and Kudo	1987	Ostensibly Domesticating
Poem E	Shellfish	Ishigaki Rin	1968	Rexroth and Atsumi	1977	Ostensibly Domesticating
Poem F	Two Billion Light-Years of Solitude	Tanikawa Shuntarō	1952	Elliott and Kawamura	1998	Ostensibly Foreignising
Poem G	No title (*Midaregami* no. 175)	Yosano Akiko	1901	Janine Beichman	2002	Ostensibly Domesticating
Poem H	instant 瞬間 shunkan	Yoshihara Sachiko	1964	James Garza	N/A (un-Published)	Ostensibly Foreignising

In the case of three of these poems, both translations have been previously published. In the case of the final poem, I will be using one previously published translation, while undertaking to produce the second one myself, with the explicit goal of employing techniques held to be foreignising by Venuti ([1995] 2008). This will be used to test the assumption that Venuti's recommendations can bridge the theory/practice divide, resulting in translations perceived as foreignising.[3]

---

[3] Against the potential charge of bias in this endeavour, I can point to the fact that it is in my interest as an early-career literary translator to produce the best translation possible, inasmuch as my name will be associated with it. I will not be using either the previously published translation or my own to predict FG. In this way I avoid preparing a stylistic analysis of my own translation.

The selection of poems was based on several criteria. The first was the availability of two different translations, one of which could be broadly described as foreignising. The second was length. Given the difficulty I expected to encounter in the sampling phase, I wanted to create a survey that could be completed relatively quickly. The third criterion was variety. As Van Peer (1980: 77) notes, "[i]t is of crucial importance for any stylistic theory that its usefulness be demonstrated in the study of a large and varied sample of literary texts." The fourth reason was unabashedly subjective, but concerned the overall reading experience: I did not want respondents to come away feeling that their time has been wasted. My hope was that readers would find something to like among the poems they received. I therefore tried to include a range of tones and themes, ranging from the light-hearted to the serious. After an extensive search, I decided upon the above texts.

I introduce the Japanese poems below. To facilitate comprehension, and to give a sense of the structure of the original Japanese, I have presented the poems with two interlinear glosses: the first gloss, in italics, gives pronunciation, while the second gloss details parts of speech.

The first poem is "Shijimi" [clam] by Ishigaki Rin (1920-2004). As Rimer and Gessel (2005: 415) note, Ishigaki worked as a bank clerk from 1934 to 1975 and "so became known as the 'bank clerk poet.'" While the travails of the workaday world certainly find expression in her work (see "Mazushī machi" [poor town] or "Kōkyō" [public]), so do the hardships of the trans-war years: her mother had died as a result of injuries sustained in the Great Kantō Earthquake, she had had three stepmothers by the time she turned eighteen, had lost two siblings by 1942, and had her home destroyed by an air raid in 1945 (Ishigaki 1998: 247-249). The war itself is explicitly addressed in several poems including "Sentō kaishi" [outbreak of war], "Aisatsu" [greeting] and "Gake" [cliff], which describes Japanese women throwing themselves off the cliffs of Saipan. Given the prevalence of such grim subject

matter in her oeuvre, one might be surprised to encounter a poem like "Shijimi":

シジミ
shijimi
clam

夜中に目をさました。
yonaka-ni   me-o   samashita
middle of night-in   eyes-ACC   opened

ゆうべ買ったシジミたちが
yūbe   katta   shijimi-tachi-ga
last night   bought   clams-NOM

台所のすみで
daidokoro-no   sumi-de
kitchen-GEN   corner-LOC

口をあけて生きていた。
kuchi-o   akete   ikite-ita
mouth-ACC   open   were living

「夜があけたら
yo-ga   ake-tara
night-NOM   grows light-when

ドレモコレモ
doremo-koremo
each and every

ミンナクッテヤル」
minna   kutte-yaru
all   eat-give

鬼ババの笑いを
onibaba-no   warai-o
demon hag-GEN   laugh-ACC

私は笑った。

<table>
<tr><td>*watashi-wa*</td><td>*waratta*</td></tr>
<tr><td>I-TOP</td><td>laughed</td></tr>
</table>

それから先は
<table>
<tr><td>*sorekara*</td><td>*saki-wa*</td></tr>
<tr><td>from then</td><td>onward-TOP</td></tr>
</table>

うっすら口をあけて
<table>
<tr><td>*ussura*</td><td>*kuchi-o*</td><td>*akete*</td></tr>
<tr><td>slightly</td><td>mouth-ACC</td><td>open</td></tr>
</table>

寝るよりほかに私の夜はなかった。
<table>
<tr><td>*neru-yori*</td><td>*hoka-ni*</td><td>*watashi-no*</td><td>*yoru-wa*</td><td>*nakatta*</td></tr>
<tr><td>sleeping-except</td><td>else</td><td>I-GEN</td><td>night-TOP</td><td>there was not</td></tr>
</table>

With respect to the third and fourth selection criteria listed above, this poem appealed to me for its idiosyncratic humour, which, along with the quirkiness of Tanikawa's poem, provides a contrast to the intensity of Yosano's and Yoshihara's poems. If Ishigaki's humour sometimes has a tinge of cruelty to it, as Ochiai (1998: 244, 246) argues, this likely has to do with the nature of the emotions expressed in her poetry. The first quality Ochiai (1998: 243-244) associates with Ishigaki's poetry is anger. This is an anger, Ochiai (ibid.) writes, that women in Japanese society are not supposed to express. On this point Ochiai (ibid.: 246) cites a passage by the American academic and novelist Carolyn G. Heilbrun (1989). Heilbrun (1989: 12-13) argues that before the publication of May Sarton's *Journal of a Solitude* (1973), in which Sarton "deliberately set out to recount the pain" she had unwittingly elided in her previous memoir, the genre of autobiographies by women had tended to "find beauty even in pain and to transform rage into spiritual acceptance." Ochiai (1998: 246) argues that Ishigaki marshalled her anger to move beyond similar kinds of "unconscious indirection" [*muishiki-no fushōjiki-sa*] and "guileless pretence" [*akuinaki fuseijitsu-sa*].

The next Japanese poem I discuss is Tanikawa Shuntarō's (b. 1931) "Nijūoku kōnen-no kodoku" [loneliness of two billion light years]. Rimer and Gessel (2005: 437) relate that since the publication of his first book of poetry in 1952, Tanikawa "has remained one of Japan's

most popular and prolific poets, and his poems also are among those most often translated into English."

This poem comes from Tanikawa's first book, which bears the same title. It is a collection whose tone is surprisingly buoyant considering the year of its publication. Kurihara (2000: 369) remarks that the poet's optimism in the aftermath of the war was a welcome change for some critics, but disappointed others: "[...] against the backdrop of lingering wartime desolation, the collection's pristine [*chōmei-na*] lyricism and sense of untouched, youthful vitality [...] seemed to signal that the future, so eagerly awaited, had finally arrived. Of course, this sense of newness, untouched by what came before, also invited the criticism that the experience of wartime did not seem to have affected the poet at all." The poem reads as follows:

二十億光年の孤独
nijūoku　　kōnen-no　　　kodoku
2 billion　　light years-GEN　　loneliness

人類は小さな球の上で
jinrui-wa　　　*chīsa-na*　　*tama-no*　　　*ue-de*
mankind-TOP　　small　　　sphere-GEN　　top-LOC

眠り起きそして働き
nemuri　　*oki*　　*soshite*　　*hataraki*
sleep　　wake　　and　　　work

ときどき火星に仲間をほしがったりする
tokidoki　　　*kasei-ni*　　*nakama-o*　　　*hoshigattari-suru*
sometimes　　Mars-DAT　　companion-ACC　　things like wanting-does

火星人は小さな球の上で
kaseijin-wa　　　*chīsa-na*　　*tama-no*　　　*ue-de*
Martians-TOP　　small　　　sphere-GEN　　top-LOC

何をしてるか　僕は知らない

nani-o   shiteru-ka   boku-wa   shiranai
what-ACC   doing-INT   I-TOP   do not know

（或はネリリし　キルルし　ハララしているか）
aruiwa   neriri-shi   kiruru-shi   harara-shiteiru-ka
perhaps   neriri-doing kiruru-doing harara-doing-INT

しかしときどき地球に仲間をほしがったりする
shikashi tokidoki   chikyū-ni   nakama-o   hoshigattari-suru
but   sometimes   Earth-DAT   companion-ACC   things like wanting-does

それはまったくたしかなことだ
sore-wa   mattaku   tashika-na   koto-da
that-TOP   entirely   certain   thing-COP

万有引力とは
banyū-inryoku-to-wa
universal gravitation-COMP-TOP

ひき合う孤独の力である
hikiau   kodoku-no   chikara-dearu
pulling each other loneliness-GEN force-COP

宇宙はひずんでいる
uchū-wa   hizunde-iru
cosmos-TOP   is distorted

それ故みんなはもとめ合う
soreyue   minna-wa   motome-au
therefore   everyone-TOP   seeks each other

宇宙はどんどん膨らんでゆく
uchū-wa   dondon   fukurande-yuku
cosmos-TOP   steadily   expanding-goes on

それ故みんなは不安である
soreyue   minna-wa   fuan-dearu
therefore   everyone-TOP   anxious-COP

二十億光年の孤独に

```
nijūoku      kōnen-no          kodoku-ni
2 billion    light years-GEN   loneliness-DAT
```

僕は思わずくしゃみをした
boku-wa   omowazu   kushami-o   shita
I-TOP   unintentionally   sneeze-ACC   did

The third poem I chose is tanka no. 145 from Yosano Akiko's (1878-1942) first collection, *Midaregami* [tangled hair], published in 1901. It is hard to overstate the impact of this collection on the early twentieth-century Japanese literary milieu. As Janine Beichman (2002a: 176-177) writes:

> Today it may be hard to imagine how daring it was to publish an entire volume by a woman poet, especially one who wrote so frankly about sexual desire. But it was not just the frankness; it was the youthfulness that struck a chord. [...] Akiko's poems were read as the words of a rebel against Tokugawa period prudery and the feudal dictates that forced the sacrifice of personal happiness to the stress on public order. [...] *Tangled Hair* became one of the central works of Japanese romanticism. It was also one of the literary sensations of the century, bringing Akiko a celebrity [...] unique among modern tanka poets [...]."

Out of consideration for my survey respondents, I have chosen a less risqué example than I might normally, but which nonetheless departs from whatever stereotypes may still linger in the public imagination about Japanese women at the turn of the last century. In tanka no. 75, the speaker describes walking into a field of flowers at night where she suspects her lover is waiting. In this poem Akiko "reverses the conventional poetic image of the woman-who-waits, the passive stationary, male-dependent female image that had dominated Japanese poetry for centuries" (Beichman 2002a: 117). A similar, but stronger effect, I believe, is to be found in no. 145. I have presented the poem vertically below to better fit the glosses in the

space provided. (Note: this is the orientation the ST would have originally taken).

夕ふるはなさけの雨よ旅の君ちか道とはで宿とりたまへ

	yū	evening
	furu-wa	what falls-TOP
	nasake-no	affection/love-GEN
	ame-yo	rain-EXC
	tabi-no	travel-GEN
	kimi	you
	chikamichi	shortcut
	towa-de	ask-not
	yado	lodging
	tori-tamae	take-IMP

The final poem I have chosen for my survey is "Shunkan" [instant] by Yoshihara Sachiko (1932-2002). It comes from her second collection, *Natsu-no haka* [summer's grave], published in 1964. The stark, repetitive piece is more than a little dirge-like, which is perhaps not surprising given the title of the collection—though perhaps it would be more appropriate to describe it as variations on a dirge, as the deaths come in a different order in each verse. While her earlier poetry often drew on bittersweet childhood memories, the tone and

vocabulary of "Shunkan" are by no means anomalous for Yoshihara. Depending on the critic, Yoshihara's list of thematic obsessions might include "confession [*kokuhaku*]," "God," "betrayal," "light," "darkness," "flesh" [*niku*], and "spirit" [*kokoro*] (Ōoka 2003: 142), or "wounds [*kizu*], pain [*itami*], tears, suffering [*kurushimi*], jealousy [and] death" (Yamamoto 1976: 159). To make sense of what part these qualities play in her work, however, it is necessary to view her output in a wider context. Kuninaka (2000: 430) provides an apt précis: this is "a poet who took the division between spirit [*kokoro*] and flesh [*niku*] as her central theme, writing in terms both elegant and merciless about the vagaries of love [*ningen-no gō-toshite-no 'ai'*]."

瞬間
shunkan
instant

海が死ぬ
umi-ga   shinu
sea-NOM   dies

けふも死ぬ
kyō-mo   shinu
today-too   dies

日が暮れる
hi-ga   kureru
sun-NOM   sets

月が死ぬ
tsuki-ga   shinu
moon-NOM   dies

けふも死ぬ
kyō-mo   shinu
today-too   dies

夜が明ける
yo-ga   akeru
night-NOM   grows light

時が死ぬ
toki-ga   shinu
time-NOM   dies

けふも死ぬ
kyō-mo   shinu
today-too   dies

人も　死ね
hito-mo   shine
people-too   die-IMP

惜しげなく
oshi-ge-naku
frugal-appearance-without

いくたび死んで
ikutabi   shinde
many times   dies

時がまたくる
toki-ga   mata   kuru
time-NOM   again   comes

死ぬ海の
shinu   umi-no
die   sea-GEN

死ぬ月の
shinu   tsuki-no
die   moon-GEN

うつくしさ
utsukushi-sa
beauty

色あせず
iro   ase-zu
color   fade-without

暮れもせず
kure-mo   sezu
grow dark-too   without

のこるなら
nokoru-nara
remains-if/when

人だけが
hito-dake-ga
people-only-NOM

醜かろう
minikukarō
ugly-SUP

人も　死ね
hito-mo   shine
people-too   die-IMP

Next, I discuss the specific hypotheses tested, and the instruments used to collect the

relevant data.

1.5. HYPOTHESES AND TEST INSTRUMENTS

My hypotheses directly address those gaps in knowledge discussed in sections 1.2.1 (The

Problem of Reception in ESL) and 1.2.2 (The Problem of Reception in TS). They can be

divided into three types: 1) those derived from the formalist and conventionalist positions on

literariness; 2) those derived from Venuti's theory of foreignisation, and 3) those postulated to

investigate whether readers process translated poetry differently than source-language poetry.

These hypotheses, which I discuss below, occur in pairs comprised of the *alternative*

hypothesis (H_a) and *null hypothesis* (H_0). The *alternative hypothesis* refers to a prediction derived from the theories being tested. In short, it "predicts that the independent variable [...] *will have an effect* on the dependent variable" (Gravetter and Wallnau 2002: 169). Meanwhile the *null hypothesis* "predicts that the independent variable [...] will have *no effect* on the dependent variable for the population" (ibid.). Per the usual methodology employed in such research, after data is collected, statistical analysis is used to determine whether the null hypothesis can be rejected.

The first alternative hypothesis (H_1) is that words and phrases appearing in a foregrounded condition, as determined by stylistic analysis, will be perceived as more striking. The null hypothesis (H_0) states that whether text is foregrounded or not will make no significant difference in the reader's identification of strikingness. If the formalist position on literariness is correct, reader variables will not significantly affect strikingness ratings.[4] The test instrument used to collect the relevant data here will be an underlining test like the one used by Van Peer (1986). I have chosen this test instrument partly because it produced Van Peer's (1986) strongest results, and partly because it is exceptionally convenient. Moreover, one would expect it to be much less involved and taxing than a computer-based protocol, the instructions to which might introduce a fatigue effect, compromising the reliability of the data. Finally, the effects of reading modality (for instance, reading on a screen as opposed to reading a physical paper) are not well understood.

The second, third, and fourth hypotheses (H_2, H_3, H_4) are related to the effects of reader variables on responses to foregrounding. They are derived from the conventionalist position on literariness, and thus are in competition with the first hypothesis. H_2 predicts that previous literary training will significantly influence the identification of strikingness. H_3 states that

---

[4] Significance here is used in the statistical sense, referring to the unlikelihood of a result occurring by chance (See Gravetter and Wallnau 2002: 182-183).

readers' attitudes toward poetry will significantly influence the same. Finally, H_4 holds that age will also significantly influence responses to foregrounding. I have chosen these variables because studies have produced inconsistent results concerning the role of training and reading habits in literary processing. Bortolussi and Dixon (1996: 472) have pointed out that research into the differences between expert and novice readers tends to attribute such differences directly to training, ignoring the possibility that they are attributable to other characteristics like "native intelligence, language skill and motivation that may not be possessed" by the students who typically serve as the novice readers in such studies. In other words, these factors might prevent them from becoming experts in the first place.

However, I am more interested in age as a potential marker of time elapsed since the conclusion of the reader's formal education. I suspect that the longer readers have been out of school, the less likely they may be to engage in the modes of reading thereby inculcated. I was motivated to investigate this by a passage in De Certeau's "Reading as Poaching" (1984: 156), in which he confronts the institutionalisation of expert readers' interpretations: "The fiction of the 'treasury' hidden in the work, a sort of strong-box full of meaning, is obviously not based on the productivity of the reader, but on the *social institution* that overdetermines his relation with the text." This includes the relationship between "teachers and pupils" (ibid.). These variables are, in short, a way of investigating the statement that the "creativity of the reader grows as the institution that controlled it declines" (ibid.: 157).[5] The null hypothesis (H_0) here states that none of the above variables will significantly influence the identification of strikingness. The test instruments I will use to collect the data here are twofold: (1) a questionnaire requesting age, highest completed level of education, and previous literary training, and (2) what is known as a Likert scale, in this case a scale of

---

[5] I am grateful to Diane Otosaka for confirming that "it" [la] refers back to "creativity" [créativité] in the French: "La créativité du lecteur croît à mesure que décroît l'institution qui la contrôlait." (Message to author, 8 September, 2020).

multiple items by which informants rate their disagreement or agreement with statements designed to measure poetry attitudes. The poetry attitude scale consists of 11 items, the majority of which were adapted or directly drawn from Van Peer's (1986: 197) poetry attitude scale. I augmented this scale with other items designed to capture a broader array of relevant information, including importance judgments ("It is important for poetry to be taught in school"), preferences ("I prefer poetry to prose") and evaluative beliefs ("I find that reading poetry enriches my life").[6]

H_5 and H_6 are the final hypotheses related to literariness. They predict that the more FG a reader indicates, the higher they will rate the text in terms of literariness (H_5) and overall enjoyment (H_6), respectively. The motivation for including these is: there seems to have been a kind of unspoken assumption in ESL that FG is a positively-valenced phenomenon—that it contributes in a necessarily positive way to the aesthetic experience of texts (including perception of literariness). Zyngier, Van Peer and Hakemulder (2007) have demonstrated that the higher the overall level of FG in a text, the more likely the reader is to perceive that text as complex. However, it was only in one of their three respondent groups that they found strong support for the claim that "readers will evaluate more highly the texts which offer more complex patterning on different levels, especially on a second reading" (ibid.: 673). H_0 states that the level of perceived foregrounding will have no significant effect on literariness and enjoyment ratings. The conventionalist position will prefer the null hypothesis. The test instrument used to collect data for H_5 will be another Likert scale on which the reader is asked to indicate disagreement or agreement with statements about the text's literariness. The test instrument used to collect data for H_6 will be the same, only readers will be asked to rate their enjoyment of the text. Fortunately, other studies have already designed items which address these variables: as Blair, Czaja and Blair (2014: 31) write, the use of questions from

---

[6] For more on these categories, see Saris and Gallhofer (2007: 40-48).

previous studies allows us to more directly compare results. To that end, I will be adapting Dixon et al.'s (1993)[7] appreciation scale: 1) 'Is this poem an example of good literature?' 2) 'Would you recommend this poem to someone else to read?' 3) 'Did you enjoy reading this poem?'

Not only has this scale been used in other studies (e.g. Hakemulder 2004; Zyngier et al. 2007), but it has the advantage of covering the three main attitudinal components in the traditional model of mind (see e.g. Rosenberg and Hovland 1960): a cognitive component, an affective component and an action tendency. However, as only one of the above questions addresses the category of literature, I will be augmenting it with another of my own: 'In your opinion, how literary is this poem?' Furthermore, I will be adding another question related to enjoyment, this one drawn from Belfi, Vessel and Starr (2018): 'How enjoyable or aesthetically appealing did you find this poem?' Asking these additional questions will allow us to further investigate the possibility that enjoyment and impression of literariness are not necessarily correlated, as is often assumed in ESL.[8]

To sum up thus far, Survey 1 will test six hypotheses related to the theory of foregrounding. Meanwhile, Survey 2 will test three hypotheses (H_7, H_8, and H_9), all of which are derived from Venuti's theory of foreignisation.

H_7 predicts that texts appearing in the 'ostensibly foreignised' condition will elicit higher scores for impression of foreignness. If the null hypothesis is true (that is, if we accept the statement that there is no relationship between foreignising translation strategies and impressions of foreignness), then a number of intriguing explanations present themselves. For instance, it may turn out that the impression of foreignness is more closely linked to reader variables like literary training than textual variables. The test instrument used to measure

---

[7] For more on this scale, see sections 2.2.2 and 5.2.6.
[8] See section 5.2.6.

impression of foreignness will be a Likert scale consisting of three questions. I discuss and justify these questions in section 3.1.3.

H$_8$ involves another underlining task: readers in Survey 2 will be tasked with identifying words or phrases that produce an impression of foreignness. The hypothesis to be tested in this case is whether readers will underline those items in the text that can be described as foreignising according to the criteria discussed in Chapter 3. It will be interesting to see whether readers primarily underline foreign vocabulary, or whether they will include other features of the texts. Inasmuch as analysis of the results will shed light on what real readers consider to be foreignising, we may discover discrepancies or gaps between what Venuti predicts and what actually obtains in empirical research.

The next hypothesis (H$_9$) derives from the assumption that translations explicitly undertaken to produce a foreignising effect can be just as well-received as translations done in an ostensibly fluent style. My motivation for including this hypothesis comes from two passages in Venuti ([1995] 2008: 121), wherein the author cites the favourable reception of two of his own translations, as well as the success of Megan Backus's translation of the Yoshimoto Banana novel *Kitchen*, as evidence that "foreignizing can cross the cultural boundaries between elite and popular readerships." H$_9$ holds that my translation of Yoshihara Sachiko's poem "Shunkan" (specifically undertaken for this project according to the criteria for foreignising translation discussed in Chapter 3) can score just as highly on literariness and enjoyment scales as the existing translation of the same text.

The final hypothesis in this project, H$_{10}$, will be tested in Survey 3. H$_{10}$ holds that readers' responses will be affected by whether the texts are explicitly presented as translations or are simply presented as 'poems.' For the purposes of this experiment, the word "differently" will be interpreted as follows: if results for the texts presented under the first condition (poem as translation) show statistically significant differences on at least one test measure when

compared to results for the other condition (poem as text), the null hypothesis will be rejected. The motivation for this experiment is to investigate the possibility of an aesthetics of translation distinct from the aesthetics of source-language texts.

The next task I set myself in this chapter is to define some key terms.

1.6. SOME KEY TERMS

First, I use the term 'literary reading' here and throughout to refer to a kind of "literacy event" (cf. Heath 1982a) in which a reader or readers engage with a text or texts belonging to the category of 'literature,' however that may be defined (by reader, instructor, institution, etc.) at the time of the reading. S.B. Heath (1982a: 93) defines a "literacy event" as "any action sequence, involving one or more persons, in which the production and/or comprehension of print plays a role." The original motivation for the creation of the term 'literacy event' was, in part, the recognition that "the narrow conceptualization of 'reading' as 'decoding the text of a book' missed a very great proportion of authentic reading […] engagements" (Gillen and Hancock 2010: 101). My adoption of Heath's term offers several theoretical affordances:

1) it allows for disagreements between people (for instance, expert readers and non-expert readers) about whether a particular text is 'literature' in a given context, thus allowing us to investigate the criteria that different readers use in applying (or not applying) the terms 'literature' or 'literary';

2) it acknowledges that readings of 'literature' take place in a variety of social settings for a variety of reasons and with a variety of consequences;

3) it does not deny that participants have engaged in literary reading just

because they have not produced an interpretation of the text;

4) it does not disqualify readers' participation in the literary system just because some interpretation they *have* produced has been deemed invalid by some arbiter;

5) it does not assume that full comprehension (e.g. being able to recognise every word in the text) is required for a literacy event to take place, as Heath's (1982b) study of early childhood reading shows;[9] and

6) it does not preclude the possibility that effects historically associated with written texts might also be produced through other kinds of aesthetic engagement (watching television shows, listening to music, etc.)

Where this last point is concerned, my adoption of Heath's term simply narrows the scope of the investigation to those texts which readers encounter primarily in written form. Thus, my approach does not deny the possibility that in certain situations the effects of literary reading may be indistinguishable from the effects of other kinds of activities. Indeed, the rise of so-called 'prestige' television over the last twenty or twenty-five years has arguably occurred not just alongside—but in conversation with—the notion that such shows resemble literature. For instance, as Bruhn and Gjelsvik (2013: 133) write, "[o]ne theme in the critical reception of [David Simon's 2002-2008 HBO series] *The Wire* is the repeated comparisons between the show and (highbrow) literature," a "theme" that has "become repeated to the degree that it lingers on the cliché […]." It is difficult to say whether such comparisons are evidence of an uptick in the description of non-printed texts as 'literary,' but it would be difficult to argue that, for instance, the awarding of the 2016 Nobel Prize in

---

[9] The idea that full literacy might not be necessary to experience literary effects is especially important as far as multilingual texts are concerned; it acknowledges that a reader of such a text may understand only one of the languages used therein, or understand them all incompletely, but that this should not automatically foreclose the possibility of literary reading. Moreover, it acknowledges that language learners may engage in literary reading despite a non-native proficiency in the particular language or language variety.

Literature to Bob Dylan did not go some way toward further legitimising the idea that the term 'literature' may encompass texts other than those primarily encountered in written form.

Viewed in this light, the term 'literary effect' must be thought of as a kind of tentative category—tentative because the set of effects that fall into this category might change from one historical moment to the next, or from one reading practice to the next. In accordance with the view of 'literature' presented here, as an emergent phenomenon produced through interaction(s) between text(s) and reader(s), I take 'literary effect' to refer to any of the changes (short- or long-term) that such interactions may have on an individual reader or readers in terms of emotion, feeling, affect, mood,[10] behaviour or cognition. If this definition seems broad, I would argue that it *must* be to allow us to investigate phenomena which are still not well understood. Whether such effects can be used to definitively distinguish literature from non-literature is a separate question. Thus, in this study, the 'literary' in 'literary effect' simply reflects that the effect was observed in the context of 'literary reading'—i.e., an engagement with a written text belonging to the category of 'literature,' however that is defined in the context of the particular 'literacy event.' The question of whose definition to accept in the case of disagreement is an interesting one, though I would argue it is the wrong question to ask here, inasmuch as it assumes that a reader's engagement with a text they do *not* consider to be literature disqualifies their responses from having any impact on the system of values that *surround* that literacy event. Indeed, the opinion that a text does not merit inclusion in a particular category (like 'literature') can be a consequential one— with a potential to alter or reinforce ideas about the meaning of that category. In this study, participants disagreed about the *extent* to which a given text belonged to the category 'literature,' as evidenced by their responses to Likert-scale questions like "Is this poem an

---

[10] For clinical distinctions between 'emotion,' 'feeling,' 'affect' and 'mood,' see e.g. Ketai (1975) and Tyng et al. (2017: 2-3).

example of good literature?" (repurposed from Dixon et al. 1993) and "In your opinion, how literary is this poem?" However, participants only rarely chose the lowest possible value in response to these questions, suggesting that most of the time, respondents considered the stimulus texts to belong either partially or fully to some category called 'literature.'

Examining category-membership as a question of degree is one way to overcome the dichotomic thinking which prevents us from seeing the internal structure of what I have called 'literary perception.' I define this as the degree to which readers perceive a given text to be like other texts which they have encountered or experienced in the past as 'literature.' (This is what I intended to capture with the adjective 'literary' in the survey question "In your opinion, how literary is this poem?") Not only does this approach allow us to recognise that non-expert readers are also capable of making subtly graded judgments which draw self-reflexively upon their own histories as readers, but it allows us (in theory) to search for increasingly subtle correlates among textual and reader variables.

Where the term 'literature' is concerned, I agree with Mari-Ann Berg's (2000: 281) assertion that it has "become increasingly problematical to use [it] […], not only because of its many historical and present senses and its general fuzziness, but also because the literary debate at large is currently fraught with a disturbing awareness that many of the values which have been taken as natural are being consciously and eagerly called in question." However, while the problematic nature of the term may seem obvious to a researcher, can it really be assumed that all respondents would be abreast of these issues? As (Widdowson 1999: 4) argues:

> […] in normal usage, a distinction tends to be drawn and signalled by
> the fact that when we are speaking of critical, theoretical or
> promotional literature, for example, we tend to put the definite article
> in front of the word: 'I'm reading the literature on…'. Whereas, when

we are referring to 'literary' writings, we leave it out, hence denoting

that some (unexplained) generic distinction has already been made: 'I

love reading literature in my spare time'/'I'm studying Literature at the

university'.

(Widdowson 1999: 4)

Widdowson enacts a further distinction by the use of the capital L in his last example. As he

explains: "'Literature' with an upper-case 'L' and inverted commas round it signifies here the

conception of that global body of literary writing which has been accredited with being—

pointedly to borrow Matthew Arnold's famous utterance—'the best that has been known and

said in the world' […]" (Widdowson 1999: 4). The important thing is that Widdowson (ibid.)

ventures that "it is the sense with the capital 'L' that is uppermost in the minds of the vast

majority of you in understanding what is now normally implied by the word 'literature'." A

lot rests upon the identity of the 'you' in the above quotation; while it is not clarified, the tone

suggests that Widdowson may be addressing the students to whom his book has likely been

assigned. Nevertheless, while he accepts Fowler's (1990: 23) point that "there is no single

linguistic criterion, or set of criteria, which distinguish(es) all the 'literary' genres from all the

'non-literary' genres," Widdowson (1999: 16) maintains that "'literature' as a concept retains

a meaningful cultural sense." In leaving the interpretation of 'literature' up to respondents, I

have hoped to cast as wide a net as possible, such that I might identify *patterns* in the

interpretation of the word. In a sense, I have supposed that both Berg and Widdowson may be

right—that 'literature' may be as much a problematic category as it is a meaningful one,

depending on which readers we ask. (This is borne out by some results of my study.)

As for my own use of the term 'literature,' unless otherwise stated, I have tried to use it

in a "neutral discursive capacity" (Widdowson 1999: 5). That is, I have tried to use it in the

awareness that it is a term whose meaning depends to a great extent on its context. If I have a

particular sense in mind for it, I will say so.

A few final remarks on key terms are in order, as they relate to the generalisability of empirical findings. While my study takes place in a research context concerned with 'literature' and 'literary reading,' my stimulus texts are all poems—a category that hardly covers the range of texts to which the terms 'literary' or 'literature' have been applied in Anglophone contexts. So: how can responses to one text type teach us about 'literary reading' *en soi*?

The answer is a complex one, and while it is deflationary regarding the generalisability of conclusions to other text types, it nevertheless allows us to recognise poetry as an *important piece* of the larger puzzle of literary reading. As Benthien, Lau and Marxsen (2019: 19) remark: "Both literariness and poeticity were used by the Formalists and continue to be used—often quasi-synonymously—in literary theory." For Shklovsky and Jakobson, and later Mukařovský, behind this quasi-synonymy was a view of literary phenomena that privileged poetry: whether the *differentia specifica* of literature was located in defamiliarisation or foregrounding, literariness tended to be associated foremost with poetry. At the same time, the concept of literature as "verbal art" (*slovesnost'*) included both poetry and "artistic prose" (Pomorska 1987: 1), such that literariness was thought to manifest in an attenuated or altered form when it came to prose.[11]

Thus, for instance, for Jakobson (1960: 374) the "prosaic variety of verbal art" occupied an uneasy theoretical middle ground, falling somewhere between "strictly poetic and strictly referential language." In an earlier essay, Jakobson ([1935] 1987: 302) seemed aware of the historical factors conditioning his poetry-centred view: "The major achievements of Russian literature in the first decades of our century belong to poetry; it is poetry that is felt to be the pure canonical voice of literature, its perfect incarnation." Nevertheless, he would famously

---

[11] See, e.g., Steiner (1984: 148).

go on to codify the "poetic function" (Jakobson 1960: 356) as the "dominant, determining function of verbal art" (see also section 2.1.1.2), in an apparently universal definition that stipulated neither a particular historical period nor language.

This poetry-centred view has informed foundational work in ESL. For instance, while Van Peer (1986) studied responses to poetry, the assumption was that FG would also be present in "longer literary works" like "prose," albeit in a "more diluted form" (ibid.: 60). However, when Van Peer describes his "standard version" of the theory of FG (ibid.: 20), it is clear that he means it to predict responses to "literary texts" (ibid.: 28) in general—not just poetry. Perhaps for this reason, at the end of his study he relates his findings back to the same broad category of literary texts (see ibid.: 182-188). (To his credit, he signals an awareness of the risks of over-generalisation [see ibid.: 177].)

In a sense, (over)generalisation can be a productive site for the generation of new hypotheses to be tested: e.g., would Van Peer's 'standard' theory be equally predictive of responses to, say, creative nonfiction? But the starting point of such research depends on *troubling* the received framework of poeticity-as-literariness; it depends on *questioning* conclusions drawn about literature from poetry. Otherwise, such generalisation might just as easily contribute to the neglect of other text types in ESL. Thus, if I speak of literary reading or response in relation to my surveys, it is with the knowledge that poetry comprises but one text type in a much larger constellation of texts that have been deemed 'literary' at one point or another in Anglophone reception scenarios. To divest from the poeticity-as-literariness model is to acknowledge the possibility that what characterises 'literary reading' for one text type might not characterise 'literary reading' for another text type.

***

In this chapter, I have discussed the motivations of the current study, as well as its potential benefits to both ESL and TS. In the next chapter, I discuss the history of formalism and conventionalism, and articulate important linkages between Russian Formalism and Venuti's theory of foreignisation. These linkages are crucial not only in justifying the decision to tackle both in the same study, but they help uncover unspoken assumptions about the translator as a source of invariance, further pointing to the need for theoretical refinement vis-à-vis foreignisation.

2. BACKGROUND

2.1. PHILOSOPHIES OF LITERARINESS

Why is it that in some studies, literary effects seem to be primarily the result of textual variables, while in others, they seem to depend more on reader variables? In this chapter I will provide everything readers need, in terms of key concepts, previous research results and methodological issues, to understand these competing claims and why they require further empirical testing. It is precisely here, in the field of ESL, where the struggle between two major philosophies of literariness—formalist and conventionalist—has been contested, and where the claims of theorists find themselves substantiated or contradicted. In assessing the strengths and weaknesses of prior research in this area, this chapter lays the groundwork for the discussion of my methodology in Chapter 3. However, we must first familiarise ourselves with the two major empirical positions on the question of literariness: formalist and conventionalist.

2.1.1. FORMALISM REVISITED

Readers familiar with the history of literary criticism may be surprised to encounter my stated concern with formalism. After all, as a school of criticism, it is mainly discussed in the past tense. When we speak of Russian Formalism, we are obliged to acknowledge its historical boundaries: it was a movement in literary scholarship which "originated in 1915-1916, had its heyday in the early 1920s and was supressed about 1930" (Erlich 1980: 11). This is not to say its ideas did not exert an enormous influence, along with those of the associated Prague School, on subsequent critical approaches. Nineteen-thirty simply marked

the death of Formalism with a capital F.

Which formalist claims did opponents single out for criticism? First, and most central to the outlook of Russian Formalism, was the notion of a specifically literary language. The working assumption of early Formalist inquiry was that the literary text—any literary text— necessarily contained textual features that "set it apart from other types of text" (Nørgaard, Montoro and Busse 2010: 24). This typological view of literariness led the Russian Formalists to concentrate, at least in the early days, on trying to formulate the defining properties of literature with particular emphasis on the linguistic basis of literary devices [*priem*].

There are several reasons why this "essentialist view," as Boase-Beier (2010: 31) terms it, no longer finds many staunch defenders, especially in its hard form presented here. The first reason has to do with our contemporary understanding of the interrelatedness of 'ordinary' or 'standard' language and 'poetic' language, as they were apt to be termed. If we consider the linguistic raw material, as it were, available to us for the purposes of making literature, we must conclude that there is nothing off-limits, no variety of speech that cannot appear in a literary context. I emphasise that I am speaking from a linguistic point of view here. Culturally, religiously—certainly legally—there may very well be proscriptions against what one can put in a book or publish in a story.

The second reason that the hard formalist position on literariness is no longer much argued is related to the first. It is the flipside of the idea that literature can be fashioned from any kind of language. As Simpson (2004: 98) puts it: "there exists no feature or pattern of language which is inherently or exclusively 'literary' in all contexts." It is crucial to note here that features from past styles of writing which we might describe as 'literary' are "more representative of specific codes or conventions of use which may change over time, rather than confirmation of the existence of a special language which in its very essence is

immutably, and for all time, 'literary'" (Simpson 2004: 98).

In some cultural contexts, specific registers or varieties may indeed be set off for types of writing we might call 'literary.' Where the STs in this project are concerned, Yosano Akiko's tanka, for instance, was written in *bungo*—a term often translated as 'literary language,' and which has functioned as a kind of catch-all for the variety of more or less prestigious *buntai* (writing styles) available to educated Japanese during the Meiji period (1868-1912).[12] Not only was *bungo* used for text-types (such as government pronouncements) far removed from the "imaginative" writing (Erlich 1980: 172)—e.g., novels, short stories, fairy tales and poetry—to which the Russian Formalists typically directed their attentions, but as Suzuki Sadami (1998: 156-157) points out, some of the genres historically written in what we now call *bungo* (e.g. tanka) were *not* always classified as *bungaku* ('literature'). Moreover, Suzuki (2013: 11) writes, "there is not a single instance of *waka* [a forerunner of tanka] or *monogatari* [Japanese tales] being referred to as *bungaku*" before the Meiji period. Such decisive changes to the reception and production of literature may not happen very often in the history of national literatures, but they *do* illustrate Fowler's (1996: 248) point that "the texts that count as literature vary from age to age."

Inasmuch as *literacy was itself prestigious* in pre-modern Japanese contexts, and far from universal, the various historical *buntai* harken back to a time when Japanese conceptions of 'literature' were more in line with 'scholarship' than with 'imaginative' writing. This has theoretical ramifications for the present study: the presence of illiteracy in reading scenarios would seem to create a niche for insider-interpreters who can control access to what in the reader-response tradition is called the 'actual text.' As Goodwyn (2013: 213) explains, the "actual text is created in a transaction between the reader and the material text and so the actual text inevitably has some personal interpretation." Where respondents' ratings for the

---

[12] See e.g. Garza (2020).

stimulus texts in this study were concerned, some readers expressed discomfort or difficulty

rating texts—it is almost as if for these particular readers, the translated-ness of the texts

called for an interpreter or translator or some other authority *beyond the one who actually*

translated the ST to further mediate their responses to the TT—to help them create a better,

more reliable 'actual text' for evaluation. In other words, it was as if they did not trust their

own personal interactions with the texts to provide a reliable basis for aesthetic judgments,

and sought to create a buffer between the self that experienced the text and the self that was

commenting upon it—to maintain a kind of plausible deniability of literary response.

This leaves the third reason why the hard formalist position on literariness is not much

argued today. While the first two reasons have been linguistic in nature, we cannot discount

the influence of the canon and other extra-textual attitudes in determining which formal

characteristics tend to predominate at any given moment. Thus, while the linguistic

characteristics of texts vary, so, too do "the circumstances of their production and reception,"

as Fowler (ibid.) notes.

But if formalism is a thing of the past, and its hardline claims more or less abandoned,

how could it possibly advance our cause? It is not Formalism as a programmatic method of

inquiry that concerns us, but rather two ideas inherited from Formalism. One is the notion

that it is the text itself that guides the act of literary reading, and thus is chiefly responsible

for the effects attested by readers. The other is more properly a theory, insofar as it seeks to

explain *how* the text incurs such attention. Moreover, it is probably the most enduring

concept to come out of the intellectual framework of Formalism, inasmuch as it is still

studied and debated today. I am referring to the theory of foregrounding.

2.1.1.1. THE THEORY OF FOREGROUNDING

For a good definition of this concept as currently mobilised in research, we may turn to Wales (2011). Her definition captures the two crucial senses of the term, which describe the external and internal relations of a literary text: in its external sense, FG is the "'throwing into relief' of the linguistic sign against the background of the norms of ordinary language" (Wales 2011: 166). Setting aside for the moment the question of what constitutes ordinary language, we can see that 'external FG' refers to the relationship between an individual text and the broader language practices in which it is situated. In the internal sense of the term, foregrounding refers to the relationship of features within the text. As Wales (2011: ibid.) writes, "[...] within the literary text itself linguistic features can themselves be foregrounded, or 'highlighted,' 'made prominent,' for special effects, against the subordinated background of the rest of the text, the new 'norm' in competition with the non-literary norm." But whether we are talking about external or internal FG, a crucial question remains: *which* norms may be expected to form the background(s) against which FG is perceived? This is a complicated issue, and one with which I grapple in section 4.1. For the time being, however, it may serve to note Pratt's (1977: 5-6) objection to the category of 'ordinary' language in Formalist scholarship:

> […] devices observed in literature were assumed to be 'literary,' to constitute 'literariness' […] because non-literature was assumed a priori not to possess the properties of literature. Hence even terminologically, the right-hand term of the poetic/nonpoetic dichotomy scarcely mattered at all. 'Nonpoetic' could be specified variously as 'practical,' 'utilitarian,' 'spoken,' 'prosaic,' 'scientific,' 'everyday,' 'communicative,' 'referential,' or any combination of these without in the least disturbing the notion of what 'poetic' was.

Whereas the Formalists supposed that poetic language was distinguished from non-

poetic language in following *additional* linguistic rules or norms, Pratt (ibid.: 10) argues that it was mistaken to assume that "only literary utterances are subject to norms other than the rules of grammar" (ibid.: 10). As she explains: "In addition to the rules of grammar (or *langue*), any utterance is subject to rules governing the use of language in the context in question" (ibid.). Indeed, there are a good many communicative contexts that the language in a poem may resemble. If we consider a poem to be a kind of "fictive discourse" in Barbara Herrnstein Smith's (1975: 773n2) conception—the term refers to "verbal structures that can be taken to *represent* natural utterances but that are not themselves" a record of the "verbal acts of real people"—then it makes sense that one potential background (BG) against which to perceive FG might be "our prior experiences" with the kind of "natural" utterance it resembles (Smith 1979: 37). But of course, each individual poem would also, presumably, be read against the BG of one's previous experience with poetry. Since we cannot feasibly investigate in depth the previous experiences of every respondent, it is indeed a theoretically fraught activity to make predictions about the kinds of language likely to stand out to respondents. This is the unavoidable element of researcher subjectivity which cannot be eliminated from such a study. However, as I endeavour to demonstrate in the methodology section, as well as in the sample stylistic analysis, decisions made on the basis of researcher subjectivity can at least be pointed out in advance and defended.

2.1.1.2 THE HISTORY OF FOREGROUNDING

The modern history of foregrounding as a literary concept begins with Viktor Shklovsky (1917). In his seminal essay "Art as Technique" ([1917] 1965: 22), Shklovsky argued in this essay that it was not the purpose of art to reflect the world but to renew our perception of it.

This was accomplished through a process he called *ostranenie*, or "defamiliarization" in Lemon and Reis's (1965) translation. In what must be the most oft-quoted passage from all of Russian Formalism, Shklovsky ([1917] 1965: 12) lays out his vision of literature as an aesthetic object:

> Habitualization devours works, clothes, furniture, one's wife, and the fear of war. [...] And art exists that one may recover the sensation of life; it exists to make one feel things, to make the stone *stony*. The purpose of art is to impart the sensation of things as they are perceived and not as they are known. The technique of art is to make objects "unfamiliar," to make forms difficult, to increase the difficulty and length of perception because the process of perception is an aesthetic end in itself and must be prolonged. *Art is a way of experiencing the artfulness of an object; the object is not important.*

Fowler (1996: 57) reads in this a "refusal to allow literary works any connection with social and historical process," but he does not remark upon the moral component of this view of literature. Its outlook is not as hermetic as later criticisms of formalism tend to suggest. Shklovsky clearly frames engagement with literature as morally improving, not just in this passage but throughout the essay, writing for example of "Tolstoy's way of pricking the conscience" (Shklovsky [1917] 1965: 13). Shklovsky may have been disposed against the idea of art as a mimetic reflection of the world, but clearly he was not uncomfortable with the idea of art commenting on the world. The stereotype of formalist inquiry as cold or dispassionate is not much in evidence here.

Nevertheless, one major criticism of Shklovsky relates to his use of the term *ostranenie*. As Van Peer (1986: 3) notes, Shklovsky was "notoriously unsystematic in his use of terms." *Ostranenie* sometimes referred to devices in the text and other times to the effects those devices had on the reader (ibid.). Yet in this bivalency, we can see the germ of the move to a

more *functional* philosophy, which is to say a philosophy in which "poetic language is not defined in terms of its properties but in terms of its function, which lies in its aesthetic effect" (Van Peer 1986: 6). This functional philosophy is more fully expressed in the works of Prague Structuralist scholars Havránek ([1932] 1964) and Mukařovský ([1932] 1964).

These scholars explicated Shklovsky's concept of defamiliarization in terms of linguistic differentiation. Whereas Shklovsky ([1917] 1965: 22) remarks that in literature "we find material obviously created to remove the automatism of perception," Havránek goes a step further, arguing that this can occur in any kind of language use. This happens when we call upon linguistic devices that are a mismatch in some way for our current "expressive purpose" (Havránek [1932] 1964: 9). When this mismatch occurs, Havránek calls the phenomenon *aktualisace*, or as Garvin (1964) introduced it to the Anglophone world in his Prague School Reader, "foregrounding."

As Pratt (1977: 24) notes, while Havránek "trie[d] to counteract the excessive rigidity of a poetic/nonpoetic dichotomy by subdividing language in a multiplicity of 'functional dialects,'" in the end he enacted the same "binary opposition" between poetic and non-poetic language that compromised the Formalist approach. Nevertheless, Havránek's paper widened the range of phenomena to which FG, as an explanatory theory, can be applied.

One significant criticism that arises from the functional view has to do with the amount or degree of FG that poetry is expected to contain. At this point in the development of the theory, the assumption seems to be that the more foregrounded the poetic utterance is, the better it achieves its aesthetic function. Mukařovský gives voice to this when he argues that the "function of poetic language consists in the maximum of foregrounding of the utterance" ([1932] 1964: 19). Mukařovský acknowledges that "the simultaneous foregrounding of all the components of a work of poetry is unthinkable" (ibid.: 20), as this would "bring all the components into the same plane and so become a new automatization" (ibid.). How then does

a work of poetry achieve its maximum of FG? As Van Peer (1986: 7) explains, this is the "result of two forces" for Mukařovský: "One of these resides in the relational character of foregrounding, the other in its consistent and systematic character." In short, "the (good) poet will avoid random deviations" while at the same time "he will try to work towards unity of the work by making the foregrounded components point in the same direction" (ibid.).

Nevertheless, where Mukařovský is concerned, in explicitly invoking the violation of norms in his conception of FG ([1932] 1964: 18), he arguably handed ammunition to critics of formalism. As numerous critics have argued, "poetry need no violate any rules of language and still remain what it is, a highly patterned and organized mode of verbal expression" (Stankiewicz 1960: 70; qtd. in Van Peer 1986: 4). However, linguistic norms are not, strictly speaking, rules—especially not in the generative grammar sense of rules, i.e. descriptions of the "subconscious set of procedures" by which native speakers "put together words" into sentences (Carnie 2013: 6). In this sense, rules pertain to "competence," or "the system of rules" acquired by speakers so that "they are able to produce and understand an indefinite number of sentences, and to recognize grammatical mistakes and ambiguities" (Crystal 2008: 92). Norms, on the other hand, may be profitably categorised as a matter of "performance," or "the kinds of language that are actually produced and heard" (Carnie 2013: 17). Moreover, since we use language in a variety of social contexts—both spoken and written—the norms associated with certain text-types may come to influence readers' expectations. We can readily find examples of poems that do not 'violate' rules. My point is: no Russian Formalist or Prague Structuralist to my knowledge ever described this sort of violation as a necessary or sufficient condition of poetry. Meanwhile, it *is* theoretically possible to describe stretches of language in any poem in terms of deviations—however slight or pronounced—from some external or internal norm.

It is not hard to imagine readers growing accustomed to particular devices, and the

effects associated with those devices subsequently becoming more and more attenuated—but this is an aspect of defamiliarisation which some of the Russian Formalists themselves (particularly Tynjanov) addressed. As Steiner (1984: 106-107) explains: "[…] a construction that appears merely 'usual' can, at one moment, become a literary fact because of the unusual nature of the immediately preceding literary tradition against whose background it is perceived, and vice versa." In other words, the "literary system […] is torn by conflicting tendencies to preserve the status quo and to change it" (ibid.: 109). In this model, literature "simultaneously contains its past and points to the future." This past is evidenced by features that *have been* effectively defamiliarising, while "its future rests with the constructions negating this automatized past" (ibid.). In its broad outlines, Tynjanov's model of literature sounds quite similar to the following assertion by Culler (1997: 40): "Literature is a paradoxical institution because to create literature is to write according to existing formulas – to produce something that looks like a sonnet or that follows the conventions of the novel— but it is also to flout those conventions, to go beyond them." Admittedly, a reader's familiarity with the theory of art as defamiliarisation might affect the way they read—*to the extent that they are conscious of it during actual acts of reading.* But if the theory explains a deep-seated cognitive process that takes place whether we like it or not, could familiarity with the description of said process really shut it down in the mind/brain? Moreover, how likely are readers to have this theory uppermost in their mind every time they read?

The work of the next major figure we discuss shows how pushing a formalist position far enough can turn it into a conventionalist one. Roman Jakobson was an important member of both the Moscow and Prague Linguistic Circles, and one of his most famous papers was a statement delivered to the 1958 Conference on Style at Indiana University and later published in Sebeok (1960). In it he distinguishes six different aspects of the communicative circuit to which all acts of "verbal communication" (ibid.: 353) are beholden. Depending on the

communicative situation, some of these factors may predominate over others, and the one that takes precedence will determine the dominant function of the communicative act (ibid.). It is not the case that communicative acts can consist of only one of Jakobson's functions; rather, different hierarchies of these functions exist in every message (ibid.). (We must be careful here to clarify what he means by "message." As Hanauer [1998: 566] explains, the notion of the *message* "refers to the structure of the utterance itself and not its content.") Where the *poetic* function is concerned, Jakobson (ibid.: 356) defines this somewhat cryptically as the "set (*Einstellung*) toward the *message* as such, focus on the message for its own sake […]." The nature of this 'focus' is never finally defined in Jakobson's essay—but a hint lies in the German word '*Einstellung*.' As Striedter (1989: 59) notes, Jakobson had in an earlier Russian-language work introduced the term *ustanovka* "as an equivalent" for the German term: "Like this German word, the Russian one can mean at once the orientation of one thing to something else, and the arrangement of all the parts within a system (corresponding to its external orientation)." But that is not all: "It can also mean the attitude of a producer or perceiver of literature" (ibid.). Thus, Jakobson's (1960: 356) use of *Einstellung* vis-à-vis the poetic function is rather convenient: it allows him to avoid pinpointing the exact source of the 'focus' on the message, sidestepping questions about the relationship between producer and receiver.

Perhaps realising that this does not definitively answer the question of what defines poetic language, Jakobson returns to the formalism he helped develop decades before. In language familiar from earlier Russian Formalist studies, he asks, "What is the empirical linguistic criterion of the poetic function? In particular, what is the indispensable feature inherent in any piece of poetry?" (Jakobson 1960: 358).

His answer is: "The poetic function projects the principle of equivalence from the axis of selection into the axis of combination" (Jakobson 1960: 358). To begin to fathom what this

means, we first need a grasp on the "principle of equivalence." Burke and Evers (2014: 35) explain it well: the principle of equivalence refers to "the patterns of language that a writer/poet chooses and employs which help to establish general connections across a text." In order to create a text in the first place, we must both select and combine words. Normally, the principle of equivalence "governs" the selection of one word over other semantically possible words (Aviram 1994: 94). For example, Jakobson (1960: 358) writes about deciding between "child, kid, youngster, [and] tot" when talking about a child. All of these are "equivalent in a certain respect" (ibid.). This occurs in both prose and poetry. As Bertens (2001: 48) writes, poetry "constantly selects items from long lists of words that are in one way or another equivalent to each other." In this respect, it is no different from any other kind of language use (ibid.).

However, unlike prose, poetry "also selects to create equivalences *between* the words it chooses" (ibid.). In Jakobson's view, poetry accomplishes this on the basis of sound values. As he writes in another essay: "The principle of similarity underlies poetry; the metrical parallelism of lines, or the phonic equivalence of rhyming words prompts the question of semantic similarity and contrast" (Jakobson [1956] 2008: 168). Thus in poetry the principle of equivalence also governs the combination of words.

But what about those poems that are not written in a particular metre, and which do not have rhyming words? It is for this reason that Jakobson (1960: 365) invokes the notion of the "intonational contour," whose design by the poet is taken to be evidence of patterning motivated by sound. Thus Jakobson can extend his theory to practitioners of free verse like Walt Whitman (ibid.).

However, if this is the "indispensable feature" supposedly inherent in all poetry (Jakobson 1960: 358), a big problem remains. It is unclear what one would point to in a text as evidence of the writer's attentiveness to 'intonational contour.' How does one know which

prosodic variables are important, and which merely incidental, in a text which one is trying to classify as poetry or non-poetry? This is an issue of what Halliday (1971: 339) has called "criteria of relevance." In short, there is "no single universally relevant norm, no one set of expectancies to which all instances" of deviation (or 'prominence' in his terminology) "may be referred" (ibid.: 341). Halliday's solution also tiptoes carefully around questions of intention: "Foregrounding, as I understand it, is prominence that is motivated. […] [A] feature that is brought into prominence will be 'foregrounded' only if it relates to the meaning of the text as a whole" (ibid.: 339). In other words, the statistical identification of 'prominence' is useful—provided that one shows the *motivation* behind it that turns it into FG. To do this, one must produce an interpretation of "what the work is about" (ibid.: 359). Not only does this shift 'motivation' from a writerly to a readerly category, but Halliday's (ibid.: 339) definition of 'relevance' assumes a kind of universality of response among readers, which in turn allows him to sidestep questions about whether the motivation identified by readers is the same as that of the writer. While this model has room for both writer and reader(s), they are effectively collapsed into the notion of function. This logic is also present in Jakobson's discussion of the 'poetic' function. Because such functions are treated as always already accomplished—i.e., the texts discussed are always already agreed to be poems—in a sense, it does not really matter what readers point to in order to determine whether the text is a poem. For this reason, Pratt (1977: 36) writes that "although the projection principle and the ideas of dominance and focus on the message *can* be profitably and appropriately used to address the question 'What makes a verbal work of art a verbal work of art?' they cannot provide an answer to the question Jakobson poses: 'What makes a verbal message a verbal work of art?'" In order for rhythm or other sound patterning to *function differently* in poetry, readers must somehow already know to pay a special kind of attention to it. In other words, readers must already know the text is a poem, and thus

approach it differently. One would not be wrong to say that this sounds like the application of additional reading *conventions* associated with poetry, i.e. it is not the text per se, but conventions that we bring to the text that change the role its features play in the process of reading. With this in mind, we may now discuss conventionalism.

2.1.2. CONVENTIONALISM

Hanauer (1998: 568) provides a succinct description of conventionalism, adapted for the purposes of empirical research from a mix of claims in the reader response tradition, drawing primarily from Jonathan Culler's *Structuralist Poetics* (1975) and Stanley Fish's *Is There a Text in This Class* (1980b):

> The conventions of poetry reading and the interpretive strategies of the reader direct the reading process. Once a text has been categorized as a poem, the reader will search for and pay attention to the textual features of the poem, in accordance with the conventions and interpretive strategies of poetry reading. The reader will not pay close attention to textual features unless the text has been categorized as a poem.

It should be noted that conventionalist theory has rarely paid attention to empirical readers. In a well-known shot across the bow of ESL, Culler (1975: 123) argued that to "take surveys of the behaviour of readers would serve little purpose, since one is interested not in performance itself but in the tacit knowledge or competence which underlies it." Broadly speaking, in place of empirical readers, reader response criticism has dealt with readers abstracted to varying degrees.

Take, for example, Culler's notion of the "competent reader" (ibid.: 121). This is a reader

whose knowledge of the "system of conventions" (ibid.: 116) guiding literary discourse will allow them to "read and interpret works in ways which we consider acceptable, in accordance with the institution of literature" (ibid.: 124). Implied in this "we" is any expert who sets out to "convince his readers that meanings or effects which he is attempting to account for are indeed appropriate ones" (ibid.). (We should note that Culler's position has since changed considerably.) The conventions alluded to above would be the ones that encourage readers to "look at the language in new ways, to make relevant properties of the language which were previously unexploited, [and] to subject the text to a different series of interpretive operations" (ibid.: 114). Chief among this particular set of conventions, Culler proposes, is "the rule of significance," which leads us to "read the poem as expressing a significant attitude to some problem concerning man and/or his relation to the universe" (ibid.: 115).

The first thing to say, however, is that his 1975 position on empirical testing is based on what appears to be a faulty analogy. Appropriately enough for someone who considers literary reading to be a "rule-governed process of producing meanings" (Culler 1975: 126), Culler modelled his notion of the competent reader on Chomsky's competent speaker, whose assimilation of a limited number of rules in the process of language acquisition allows them to produce an indefinite number of 'well-formed' utterances which "other speakers can understand immediately" (Chomsky 1964a: 50).

To explain this phenomenon, Chomsky (ibid.: 51) proposes the notion of competence: "On the basis of a limited experience with the data of speech, each normal human has developed for himself a thorough competence in his native language." This can be thought of as a finite "system of rules" which is capable of generating an infinite number of "well-formed sentences" (ibid.). Crystal (2008: 219) remarks that "well formed" [sic] is an "alternative term for 'grammatical […].'" In other words, a "sentence is well formed if it can be generated by the rules of a grammar" (ibid.: 520).

In Chomsky's model, competence is contrasted with performance. While competence predicts what *can* be produced (for instance, a sentence that goes on forever, joining clause after clause after clause...), performance refers to what actually gets produced. Pius ten Hacken (2007: 44) usefully sums up the difference thus: "although competence underlies performance, performance is not a straightforward reflection of competence." In other words, performance is "filtered and sometimes distorted by various cognitive constraints" (ibid.: 53). Thus, in the Chomskyan view, the study of performance would yield only a partial understanding of the phenomenon of human language; what is called for is the study of competence.

Culler's argument for literary competence proceeds along similar lines. Empirical testing would produce only data about performance, which after all, Culler (1975: 123) argues, "can be influenced by a host of irrelevant factors [...]." This analogy between Chomskyan competence and literary competence falls apart at several different levels. Firstly, acquisition of one's first language is a natural phenomenon. We may spend a considerable amount of time in school learning how to write, but writing is a secondary system, a technology which allows us (among other things) to systematically represent speech sounds. Teachers may 'correct' our speech, but this may be better understood as an effort to bring one naturally acquired variety of speech in line with a more institutionally-approved variety. Moreover, teachers may give us advice on how to better express our ideas, but this is rather socially acquired knowledge "about how to make communication successful" (ten Hacken 2007: 53).

The idea at the crux of Culler's analogy, namely that one's understanding of both language and literature "depends on mastery of a system" (1975: 114), is further complicated by the fact that first-language acquisition happens "subconsciously" (Tsujimura 2014: 1) and "without conscious effort" (ten Hacken 2007: 73). Moreover, there seems to be a natural endpoint to acquisition of first-language competence. As Grimshaw et al. (1998: 237) note:

"Lenneberg (1967) formally proposed a critical period for language acquisition that extends from infancy until puberty." By contrast, it is unclear at what point, and after how much effort and training, a reader becomes a competent reader in Culler's scheme.

Given that the conventions of texts and genres do not correspond to structures in the brain as grammars do in Chomsky's theory, is it not more reasonable to suppose that there are gradations of knowledge of literary conventions, even among 'experts' and 'competent' readers? If it is true that the writing of literature is "made possible by the existence of the genre" which informs an author's choices (Culler 1975: 116), then it seems possible that one could continue to study ever more subtle differences between texts in a genre, differences which may be mobilised as a kind of symbol of one's superior familiarity with conventions in order to invalidate or call into question other competing interpretations or classifications.

Culler (1975) is aware of the challenge that disparate interpretations pose to his theory. He writes that the fact that "critics should differ so widely in their interpretations might seem to undermine any notion of a general literary competence" (ibid.: 122). To this objection he answers that the "model does not imply there must be unanimity on any particular count" (ibid.: 122-123). Rather, for Culler, the true task of the theory is to explain "how it is that a work can have a variety of meanings but not just any meaning whatsoever [...]" (ibid.: 122). That is, the purpose of the theory is to "account for the range of acceptable meanings which works can have for readers of literature" (ibid.: 120).

However, is it not the case that 'unacceptable' interpretations come from competent readers, too? The notion of acceptability is one of several guises in which we encounter the conventionalist preoccupation with constraints on response. Given that the common charge against reader response criticism was that it "promote[d] radical relativism and indeterminacy, leaving meaning to the subjective whim of each individual reader" (Leitch et al. 2010: 1522), this is perhaps an understandable preoccupation to have.

Nevertheless, in positing the idea of a competent reader—essentially a kind of 'native speaker' of literature, whose acquisition of a 'grammar' for reading literature is all that is required to be able to produce "acceptable" (ibid.: 124) interpretations, in the same way that native speakers are able to produce well-formed sentences—Culler (1975) strains to position literary competence as a phenomenon that does not require empirical confirmation. Yet as I have tried to demonstrate above, the notion of literary competence modelled on linguistic competence is flawed from the start. A more appropriate analogy for literary competence might be pragmatic competence. Chomsky (1980: 224) has defined this as the "knowledge of conditions and manner of appropriate use, in conformity with various purposes." In theory, this split between grammatical and pragmatic competence allows for a "person to have full grammatical competence and no pragmatic competence" (ibid.: 59). Pragmatic or discourse competence would seem to be rather more differentially distributed than grammatical competence among people belonging to the same language community. After all, the acquisition of discourse competence is premised on the "social nature of literacy development" (Gutierrez 1995: 34), and we know that not everyone has the same opportunities to be exposed to certain kinds of discourses, genres or text types.

In summation, when we take the view of literary competence as akin to pragmatic competence—i.e., an unevenly distributed rather than categorical phenomenon—the usefulness of empirical testing emerges quite clearly. The differences in response which surveys can attest should not be immediately written off as lapses to do with suboptimal performance. Rather, when we can sensibly rule out the possibility of interference from other factors, diversity of response may indicate meaningful differences in one or more reader background variables.

2.2. THE RESEARCH

So far we have covered both the formalist and conventionalist positions of the question of literariness. However, this discussion has taken place mainly in the realm of theory. In this section I will discuss the important findings of the major empirical studies in the field. Then I will discuss certain methodological issues that may have influenced results. Finally, I will describe the methodological safeguards I put in place to help my project avoid the same pitfalls.

2.2.1. INCONSISTENT RESULTS

Empirical research on reader response has yielded inconsistent results concerning the role of reader background in the perception and appreciation of FG. As Miall (2006: 18) notes, the "recent position of empirical scholars has been against" the position that FG is a "defining characteristic of literature"; however, the "agreement" attested in responses to FG "appears to be due principally to the text" (ibid.: 19). Indeed, some studies, like Van Peer (1986), Hoffstaedter (1987) and Miall and Kuiken (1994), have shown considerable agreement among readers in the identification of foregrounded elements, despite differences in the levels of literary training of respondents. On the other hand, studies like Dixon et al. (1993), Hakemulder (2004) and Zyngier, Van Peer, and Hakemulder (2007) have produced evidence suggesting that the appreciation of more complex texts (whose complexity is measured in terms of FG) depends on literary socialisation. As we will see, these discrepancies are partly explained by differences in methodology, and partly by the way researchers have interpreted their own results.

One of the earliest studies on the nature of literary response was carried out by I. A. Richards (1929). Richards' interest in personal responses to literature came not only from an academic interest in psychology and linguistics, as Goodwyn and Findlay (2003: 101) note, but from a "revulsion against the insidious and oppressive effects of wartime propaganda" (ibid.). As part of a course he delivered in 1925, 1927 and 1928, Richards distributed sets of four poems at a time to his students, the majority of whom were "undergraduates reading English with a view to an Honours Degree" (Richards 1929: 4), though West (2002: 208) mentions that also present during the 1925 course were "a few interested others, including T.S. Eliot, F.R. Leavis and Mansfield Forbes." The students were then given a week to "comment freely in writing upon them," at which point responses were collected, compared, and incorporated into Richards' lectures (Richards 1929: 3-4). Richards remarks that he was careful not only to avoid predisposing students "for or against any poem," but also to preserve respondents' anonymity, for "only through anonymity could complete liberty to express their genuine opinions be secured for the writers" (ibid.: 3). Richards did not reveal the authorship of the poems beforehand, and participation was not compulsory.

In short, Richards found that reactions to the poems "were highly variable and idiosyncratic" (Peplow and Carter 2014: 441). They varied so widely, in fact, that Richards (1929: 12) expressed difficulties in organising them: "The astonishing variety of human responses makes irksome any too systematic scheme for arranging these extracts. [...] I shall proceed poem by poem, allowing the internal drama latent in every clash of opinion, of taste or temperament to guide the arrangement." Opinion, taste, temperament: all of this supports the notion of reader background as a primary factor in literary response. At the same time, he attributes some differences to what we might call 'discourse competence.' Among the chief

difficulties he describes is the frequent lack of comprehension of the propositional content of the texts (ibid.: 13). As he writes, "The most disturbing and impressive fact brought out by this experiment is that a large proportion of average-to-good (and in some cases, certainly, devoted) readers of poetry frequently and repeatedly […] fail to make out its prose sense […]" (ibid.). This is evidence against Culler's (1975) notion of 'literary competence': if even good or devoted readers of poetry sometimes struggle to comprehend the language of texts, this would seem to be less an issue of their familiarity with the conventions for reading poetry in general, and more of an issue of their familiarity with *particular poetic discourses*.

Peplow and Carter (2014: 441) describe the experiment as "a little crudely designed by modern standards," and admit that Richards' analysis of his data "may be questionable" (ibid.). For one thing, a variety in response would not be the least surprising in a class in which production of original insights was emphasised. (I cannot speak to Richards' emphases as an instructor, but I would be surprised if he discouraged originality of thought.) A second danger in using a written protocol is, as Bortolussi and Dixon (1996: 476) relate, that it does not isolate interpretation from "the mental processes required to generate extended verbal responses." In other words, it is "important to ensure that what is assessed is truly the subjects' reception of a given text, and not their mastery of the critical or rhetorical language of literary criticism" (ibid.).

The next study I will discuss supports what its author calls a "modified Conventionalist [sic] position, in which both textual features and conventions of reading play a role in the reading of poetry" (Hanauer 1998: 565). Hanauer's goal was to arbitrate between the claims of three groups of theorists, referred to in his schema as Formalist, Stylistic and Conventionalist.

According to Hanauer (ibid.: 567), the Stylistic position is represented by scholars like Thomas Sebeok, Roger Fowler and Ronald Carter; like the Formalist position, it maintains

that the "textual features of the poem direct [its] reading [...]." Thus he groups the Formalist and Stylistic groups together. Hanauer (1998) derives two readily testable predictions from the Formalist/Stylistic and Conventionalist positions. The Formalist/Stylistic position predicts that the reader, when faced with a text that has the "characteristic textual features of a poem," will be "forced to pay close attention to the specific textual features of the poem and use them to interpret the poem" (Hanauer 1998: 569). Meanwhile, the Conventionalist position predicts that once a text "has been categorized as a poem, the reader will search for and pay attention to the textual features of the poem, in accordance with the conventions and interpretive strategies of poetry reading" (ibid.). In other words, the reader will pay less attention to such textual features when the text has not specifically been categorized as poem (ibid.).

To test these predictions, Hanauer had 80 native English-speaking students read one of four versions of the same poem, presented under one of two conditions. The poem in question was "The Twilight Turns from Amethyst" by James Joyce. His aim was to measure attention capture through a verbatim recall test, which took place after subjects had read the text twice and answered a comprehension question. Graphic and phonetic features were altered to yield four different versions of the poem, each distinguished by a different combination of formal features.

Hanauer accomplished this by rearranging half of the poems so that they resembled a prose paragraph (the original poem consisted of three stanzas, each with four lines). Half of the resulting prose-formatted poems, as well as half of the stanzaic poems, were then rewritten to reduce "the frequency and type of sound patterns in the poem" (Hanauer 1998: 570). In practical terms, this meant replacing a number of the poem's original words with semantically near alternatives to eliminate end rhyme and alliteration (ibid.: 570-572). These four versions were presented to students in one of two conditions: in the first condition, the text was referred to as "the POEM" [sic], while in the second condition the text was simply

called "the text" (ibid.: 572-574). After reading the text twice, students were given an additional copy of their particular text with five words deleted. Their task was to recall the deleted words.

Hanauer's data showed that under both conditions (poem as "the POEM" and poem as "the text"), students' recall was better for the stanzaic versions. In other words, the "level of phonetic information" did not seem to have "affected the frequency of recall [...]" (ibid.: 573). This was rather surprising, inasmuch as the patterning of phonetic information is a core feature of formalist explanations of poetry. Equally surprising was that frequency of recall was actually slightly higher for poems presented in the instructions as "the text," although Hanauer states that a subsequent analysis of variance test showed this to be statistically insignificant (ibid.: 574).

These results need unpacking if we are to understand why Hanauer says they support a "modified Conventionalist" position (ibid.: 565). After all, the text with the lowest rate of recall had indeed been described to respondents as a poem, and retained all original phonetic features—the only difference was it had been presented in a prose layout. This would seem to go against the conventionalist claim that external categorisation affects how we read poetry; since it had been described to readers as a poem, it should have scored higher if the conventionalist position is correct. However, the formalist position also falls down here; if it is indeed 'poetic language' that captures the most attention, this version of the poem should have scored higher recall rates than, for instance, the prose-formatted, phonetically altered version. But it did not. As I mentioned above, phonetic features had no effect on recall for *any* of the versions that students read.

The tentative conclusion Hanauer (1998: 577) makes is that perhaps the visual outlook of a poem should be interpreted as a convention, rather than a formal feature. While graphic features of poetry have traditionally been associated with the Formalist/Stylistic camp, he

cites a previous study of his own which shows that the graphic outlook of a poem is "central" for its categorization as such (see Hanauer 1996). Therefore, it is possible that "the graphic form announces that the text before the reader is a poem and should be read according to the conventions of poetry reading" (Hanauer 1998: 577). In support of this shifting of graphic features from formalism over to conventionalism, he goes on to cite Culler (1975) on the role of typography in helping to establish expectations (Hanauer 1998: 577). But if stanzaic organisation is something which, unlike the announcement that a text is a poem, actually inheres in the text (i.e. is detectable by the application of a rule), is it really correct to dismiss it from the category of formal features?

One methodological issue deserves mention: his use of a recall test as a parameter of attention capture. Van Peer (1986) used a similar test as a parameter of foregrounding, but found that in some cases "items from the BG [background] were recalled with greater frequency than the FG items" (1986: 53), which led him to wonder if the results from his recall test had ultimately been contaminated by other variables. One of these variables is the fact that "concrete nouns are generally recalled more easily than abstract ones" (ibid.: 99). Thus, even though the words Hanauer deleted "were chosen because of their phonetic and graphic salience" in the poem's original version (Hanauer 1998: 572), some of these words may have been more difficult to recall in and of themselves. (It is unclear in the paper itself which words he deleted from the phonetically altered versions.)

The next paper I discuss also tests claims from both text-oriented and reader-oriented approaches. Dixon et al. (1993) hypothesise that "depth of appreciation," or the increase in readers' appreciation of a text from first to second reading, will be "larger for literary texts than for other text types" (Dixon et al. 1993: 17). This type of effect is known as an "emergent effect" because it is "produced later, after the initial reading" (ibid.: 14). In contrast to earlier studies like Van Peer (1986), who studied more or less spontaneous

reactions to literature, the researchers approach the question of literariness through effects that "emerge over time" (Dixon et al. 1993: 14). At the same time, they are careful to point out that not all emergent effects are literary, and not all literary effects are emergent (ibid.).

The goal of their research is to show whether texts that generate different interpretations for different groups of readers nevertheless generate a comparable amount of literary effects for both groups. Because each group may experience a different set of literary effects in forming their interpretations, the researchers implement a special measure that "provides a global index" of these effects, no matter what they may be for each group (ibid.: 17). This is the "depth of appreciation" measure, which "reflects an aggregate of all the emergent effects of the text" (ibid.). This measure consists of three questions: "Is this story an example of good literature?" "Did you enjoy reading this story?" and "Would you recommend this story to someone else to read?" (Dixon et al. 1993: 19).

However, since the researchers also want to judge whether "*particular* effects [...] are common in the population" (ibid.), they recognise the need for another measure. This additional measure will gauge the so-called "potency" of the interpretation, a term which refers to how much a given effect is involved in the interpretations shared by each group of readers (ibid.: 15). The researchers devote one experiment to depth of appreciation and another to potency.

For their first experiment, the researchers secured the participation of 45 psychology undergraduates, and had them read one of two texts: one was a translation of the Jorge Luis Borges story "Emma Zunz," and the other was a story entitled "Death Was Her Dowry," taken from what the researchers describe as a "true detective" magazine (ibid.: 18). As the researchers remark, the stories were "superficially similar in that both concerned a murder and involved a woman as a central character" (ibid.: 18). Students were instructed to read their story once, then answer a set of questions designed to gauge enjoyment. After this, they

were asked to reread their story and respond again to the same questions.

Results showed that appreciation for both stories improved upon a second reading, but that appreciation for "Emma Zunz" improved "substantially more" (ibid.: 19). Interestingly, the Borges text was initially rated lower than the crime magazine story. However, since depth of appreciation was more than four times higher for the Borges story, the researchers conclude that "Emma Zunz" is perceived as more literary by this population (ibid.: 21).

However, they do not discuss the implications of the fact that the final average appreciation levels of the two stories are not that far apart: "Emma Zunz" tops out at 13.41 while "Death Was Her Dowry" is not far behind at 12.87 (ibid.: 19-20). Even if appreciation for the Borges story improved the most, it seems respondents ultimately enjoyed both stories to roughly the same degree. In a sense, it is fortunate for the researchers that they are using depth of appreciation as their measure; otherwise, they might have to conclude that both stories were perceived as literary.

The results from the researchers' second experiment seem somewhat sturdier. For this experiment, they enlisted 48 more psychology students. None of these students had participated in the first experiment. Whereas in the first experiment the researchers were interested in depth of appreciation, this time they wanted to test whether "narratorial ambiguity" (Dixon et al. 1993: 24) was the main effect influencing reader evaluations. The choice of this feature was motivated by the researchers' own interpretation of the theme of "Emma Zunz."

Respondents were divided into two groups based on self-reported levels of reading for pleasure: one group of 24 was comprised of "frequent" readers, with the other group of 24 being "infrequent" readers (ibid.: 25). In this experiment, readers in both groups received either a modified version of "Emma Zunz" or a modified version of "Death Was Her Dowry." The modified version of "Emma Zunz" was rewritten in such a way as to remove the

narratorial uncertainty which was central to the researchers' own interpretation of the story. Meanwhile, "Death Was Her Dowry" was "changed to include precisely the same cues to narratorial ambiguity that were removed" from the Borges text (ibid.: 25).

After obtaining results, the researchers turned to the depth of appreciation scores from the first experiment. Respondents from that test were likewise divided into frequent and infrequent readers. After comparing results from both experiments, the researchers were able to conclude that altering "Emma Zunz" in this way knocked depth of appreciation scores almost down to zero for frequent readers. In other words, whereas frequent readers in the first test appreciated the unaltered "Emma Zunz" rather more the second time, frequent readers in the second test appreciated the altered version only a little more. "Thus," the researchers write (ibid.: 27), "it would seem that for these readers narratorial ambiguity in *Emma Zunz* [sic] is an important feature, and that without it many of the emergent effects that contribute to depth of appreciation in the original are lost." Furthermore, adding ambiguity to "Death Was Her Dowry" actually brought depth of appreciation for this group below zero; in other words, whereas frequent readers in the first experiment reported a slightly higher preference the second time they read "Death Was Her Dowry," frequent readers in the second experiment seemed to like it a little less after reading the 'ambiguous' version twice (see Fig. 2 in Dixon et al. 1993: 27).

Where infrequent readers were concerned, the researchers report that these readers were not affected in their depth of appreciation scores for either story (ibid.: 28). The researchers point out that this group showed depth of appreciation for both versions of "Emma Zunz," concluding that "whatever interpretation contributed to their superior evaluation of this story on second reading, it probably did not include aspects of the narrator's knowledge and language" (ibid.).

These results suggest that frequent readers agreed with the researchers' view of the

importance of narratorial ambiguity in the enjoyment of the Borges text. The presence or absence of this feature was a predictor of whether or not such readers would enjoy the text more on second reading. It should be noted that the purpose of this study was not to test whether textual variables or reader variables principally direct literary processing. On the contrary, the test results demonstrate their "interaction" (ibid.: 31). As the researchers write: "Our method provides a viable alternative to the polarized 'text-bound' versus 'reader-oriented' theories of literary reception and interpretation. It allows one to identify objective properties of the text that affect literary processing (text features) while also giving due weight to the contribution of the reader to the appreciation of literature and literariness (reader features)."

I agree with the authors' statement that their "framework motivates a much needed focus on the interaction between text features and reader characteristics" (ibid.). However, the study does not, in my view, provide strong support to the notion that "depth of appreciation is higher in literature than in non-literature," as its authors conclude (ibid.). Firstly, the results were based on only two stories. Secondly, the emphasis on emergent effects is based on the assumption that the literary effects that *really* count "are generated only later through study and reflection" (ibid.: 14), while it is possible—perhaps even likely—that readers would form an opinion of the 'literariness' of the text at a much earlier point. Finally, if, as the researchers propose, a text is "literary if it generates a large number of (common) literary effects in a population" (ibid.: 14), who decides what constitutes a large number?

The next study I discuss is Emmott, Sanford and Morrow (2006). In this study of attention processing during prose reading, the researchers were interested in whether text fragmentation, in the form of sentence fragments and mini-paragraphs, can "change readers' alertness to details in text" (Emmott, Sanford and Morrow 2006: 2). They cite several reasons why they undertook this project. One was to produce empirical data about "fundamental

stylistic assumptions about foregrounding" that have yet to be tested (ibid.). This would act as a corrective to text-based approaches which "take for granted that linguistic norm-breaking has an effect on readers" (ibid.). Another reason is to demonstrate the feasibility of research that combines humanities and social sciences concerns—specifically, those of psychology (ibid.). Finally, they explain that "there seems to have been relatively little sustained study of unusual paragraph breaks" (ibid.: 4).

The researchers posit several hypotheses about the narratological and rhetorical effects that such devices have on readers, including their use in emphasising "plot-crucial information for the reader" (ibid.: 10). This leads to an interesting discussion about what the New Critics referred to as the intentional fallacy. This refers to the idea that the "design or intention of the author is neither available nor desirable as a standard for judging the success of a work of literary art" (Wimsatt and Beardsley 1946: 468). None of the studies discussed thus far have much grappled with this issue. It is refreshing, therefore, to see Emmott, Sanford and Morrow (2006) approach the question of intentionality from an interesting angle. They justify their investigation of potentially unintended effects by positing a special category for them:

> The fact that writers may not always be using emphasis as a deliberate stylistic device does not [...] undermine the psychological investigation of fragmentation as a potential attention-capturing device. Even in cases where providing emphasis does not seem to be the main intention of the writer, psychological methods can still be used to test whether the fragmentation has the secondary effect of making information in the fragments more noticeable to readers and/or more memorable. This might be useful for applications such as testing the readability of texts in literacy research.

This view accords with Saldanha's (2011a: 30) argument that linguistic features need not be dismissed simply because they are not intentional: "Stylistic patterns may have a clear function in the sense that we can see a point, often an aesthetic point, to their prominence, while not being necessarily a conscious strategy on the part of the text producer to create one particular meaning." I have been careful in my own analyses not to ascribe effects to the author's or translator's intentions—a category which I believe is unknowable in any fully conclusive sense. No one has ever directly observed an intention; the best we can do is make the case that *something else*—an additional piece of commentary, a remark in an interview, etc.—accurately reflects an intention. Despite not appealing to intention, the work of predicting areas in a text that might capture special attention from readers can indeed proceed; one does not need to know the intention of a person who suddenly shouts 'fire' in a crowded cinema to predict that the utterance would probably attract attention.

To begin their experiment, the authors prepared 36 short passages incorporating stylistic features from various styles of writing, including "literary texts, popular fiction, (auto)biographies and narrative embedded in popular science writing" (Emmott, Sanford and Morrow 2006: 20). They furthermore decided that each text should appear in three versions. In each version, the crucial text fragment or very short sentence would appear in a different place. In one version, the segment would be "assimilated into an earlier sentence" (ibid.: 21). In another version, the segment would be presented as its own fragment or sentence (but not in its own paragraph). And in the third version, it would be presented as its own mini-paragraph.

The researchers recruited 24 undergraduates to take part. Eight were given the first computer file, eight received the second file, and the final eight received the third. The students were instructed to read each passage once, then press a button which would call up

the same passage again, but with one word from the critical segment potentially changed (there was a roughly fifty-fifty chance of this happening). The students were instructed to notify the experimenter if they noticed a change, and to explain what they thought it was.

The results of the experiment revealed the following: readers were more likely to notice changes when fragments or very short sentences occurred towards the end of the short texts (i.e. in conditions two and three), than when the same items were assimilated into the preceding sentence (i.e. condition one). These findings support the idea that "more careful, deeper, processing is occurring for short stand-alone items" (ibid.: 23), as the researchers expected. However, as they also point out, it makes "no reliable difference" whether that short stand-alone item gets its own paragraph (ibid.: 22). This was contrary to expectation.

The researchers suggest that "simply putting the text into its own sentence fragment or very short sentence may be enough to maximize attention" (ibid.: 23). This would explain why further singling it out by giving it its own paragraph did not seem to make a difference. In any case, this is evidence that textual features affect reader behaviour, at least in terms of "depth of processing" (ibid.: 1). This concept, which the researchers have borrowed from psychology, holds that "if we pay attention to a particular word, then its meaning will play a greater role in our ongoing interpretations" (ibid.: 18). (It must be noted here that when the researchers describe readers as forming an 'interpretation,' they seem to mean forming a "full semantic representation" of the text [ibid.: 4], i.e. making logical connections between the various parts of the text.)

While I have said this is evidence in favour of a text-directed model of reading, we must be careful in interpreting it as supporting the formalist position. In the first place, results for the "literary texts" (ibid.: 3) are not compared with the results for any of the other text types. As it stands, their results seem to support the idea that narrative texts of all types (whether or not they are "literary" or "popular" [ibid.: 3]) seem to operate by generally the same

mechanism of attention capture.

The next study I would like to discuss is Zyngier, Van Peer, and Hakemulder (2007). This study investigates the role of complexity, as opposed to novelty, in readers' evaluations of literariness. Here, the researchers propose that it is not FG per se that directly accounts for depth of appreciation from one reading of a text to the next, but rather textual complexity as achieved through FG. Adopting the mathematical view that beauty (in nature, at least) arises from symmetry (Zyngier et al. 2007: 655), the researchers theorise after Reber et al. (2004: 365) that "beauty is grounded in the processing experiences of the perceiver that emerge from the interaction of stimulus properties and perceivers' cognitive and affective processes." The researchers refer to this as the "interactionist perspective" (Zyngier et al. 2007: 655), a term which reminds us of the description of text variables and reader variables as an "interaction" in Dixon et al. (1993: 31).

In approaching the literary text as something that belongs to a much wider group of stimulus objects, a group that could include, for example, butterflies or the Sagrada Familia (inasmuch as these are all apprehended by the visual system), the researchers situate literary processing within a much more fundamental theory of perception rooted in biology. Van Peer (1986: 21) has written that the concept of FG comes "from an analogy with a fundamental characteristic of human perception, i.e. the necessity to distinguish, in the act of perceiving, a figure against a ground." In terms of the papers discussed here, one has to go all the way back to Shklovsky (1917) to find literary processing linked to any other theory of aesthetics or perception. The study by Zyngier et al. (2007) is thus unique in trying to understand the phenomenon of literary reading in terms of how we perceive objects more generally.

The researchers hypothesise that the more FG readers perceive in a text, the more likely they are to rate it as complex. At the same time, the researchers are careful to point out that they do not believe FG and complexity are a guarantee of "literary and aesthetic quality"

(Zyngier et al. 2007: 659-660). But they do predict that "texts that offer more foregrounding and are more complex to the reader may be evaluated higher on a second reading, as they tend to prolong the reader's experience" (ibid.: 660).

To test their hypothesis, the researchers decided to focus on a specific stylistic feature as an index of foregrounding, namely, free indirect discourse (FID). The authors cite Fowler's (1989: 79; qtd. in Zyngier et al. 2007: 660) definition of FID as "the musical polyphony [which] refers to the co-presence of independent but interconnected voices." The researchers chose FID because it is not exclusive to English and because of its extensive use in "modern literary narrative writing" (Zyngier et al. 2007: 661).

The researchers admit that neither FID nor FG are "exclusive to literary texts" (ibid.: 661-662). Instead, they hypothesise that such features may contribute to the impression that a text is more complex, and that the 'depth of appreciation' scale first used by Dixon et al. (1993) will pick this up. The researchers seem to want us to infer that the literary and non-literary texts may be distinguished along these "functional" (Zyngier et al. 2007: 662) rather than strictly formal lines.

For their textual stimuli, the researchers used three texts of varying complexity. First, each researcher independently rated thirty passages of British and American fiction for degree of complexity. The scale included "phonological, morphological, syntactical, and discoursal features and levels of possible meaning" (ibid.: 665). Next the researchers compared their ratings. The texts which showed the highest agreement were selected for the experiment. The texts they decided on had a number of features in common: they were all narrated in the third-person, had a female protagonist, represented that character's feelings through FID, and were on the theme of love (Zyngier et al. 2007: 664). Apart from being excerpted from novels, the texts were not altered.

The researchers secured the participation of 115 participants from "three different

cultures" (ibid.: 668). The first group consisted of 43 university students of English and teachers of English in Brazil. An additional 47 were third- and fourth-year students of English at a university in Cairo. And a final 25 were literary theory students at the University of Utrecht. Each student was given a questionnaire containing the three texts and the 'depth of appreciation' scale for each text. An additional rating scale of affective qualities ("boring; complex; deep; intense; powerful; rich; senseless; striking; tiresome; trivial; unimportant; weak") was also included for readers to fill out (ibid.: 668). Readers were instructed to read each text once and fill out the questionnaire. These were then submitted to the researchers. Next, they were instructed to read each text again and fill out an identical questionnaire. At this point, information on reader background was also collected, including reading habits (ibid.).

Results showed that respondents agreed with the researchers' judgments of textual complexity. While the researchers' interpretation of their results is that the level of FG of each text has predicted the level of complexity perceived by readers, it is not clear that readers' evaluations of complexity are actually due to perception of FG. Given the background of the readers as non-native speakers, it could very well be that ratings for complexity simply reflect levels of difficulty involved in comprehension.

What is interesting here, given Van Peer's (1986) and Miall and Kuiken's (1994) earlier results, is that complexity does not correlate with depth of appreciation for all three groups. In fact, it was only the Dutch group that reported significant differences between first and second readings of the texts. Results for this group bore out researchers' predictions that the least complex text would be appreciated less on a second reading, while the most complex text would be rated higher, and the text in the middle would see little or no change (Zyngier et al. 2007: 670).

Meanwhile, the Brazilian group and Egyptian group on average were not very

enthusiastic about any of the texts. Neither group attested high depth of appreciation scores. With regard to the Brazilian readers, the researchers report that, based on responses to the affective scales, none of the texts were "very attractive to those readers, neither on a first nor on a second reading" (ibid.: 671). The Egyptian readers showed a "highly similar pattern" to the Brazilians (ibid.: 672). In short, the Dutch students were the only group whose responses went consistently in the predicted direction—the more complex the text, the higher the depth of appreciation score, and the higher the affective ratings.

Surprisingly, the researchers interpret their findings as supporting the formalist position. Why should this be the case, if only the results from the Dutch group support the 'depth of appreciation' model from Dixon et al. (1993)? The researchers write that "it is necessary to analyze the data from a different perspective" (Zyngier et al. 2007: 675). As for why they "found good overall support for our theoretical conjectures in the Dutch group" but "little or no support in the Brazilian and Egyptian groups," they argue that "such predictions hold especially for populations of frequent readers" (ibid.). While it is true that the Dutch students reported the most time spent per week in leisure reading, the problem with this explanation is that, by invoking the variable of reading habits, the researchers are inadvertently making the case for a conventionalist, or at least interactionist, interpretation. After all, reading habits are a reader-background variable.

Moreover, while the difference in leisure reading between the Dutch and Egyptian students was pronounced (9.95 hours per week versus 3.9), the difference between the Dutch and Brazilian students was not all that great (the Brazilians reported 8.6 hours per week). This suggests that appreciation may not simply be a matter of the *quantity* of literature one reads, but the *kind* of literature one reads. Again the issue of the canon rears its head: can we really assume that readers from different countries with different literary traditions value the same textual qualities?

Thus, the researchers' explanation seems a fairly drastic leap, especially since the differences might be explained by the data they already have. The Dutch students were moreover students of literary theory, which might explain why their responses alone conformed to those predicted by expert readers. Finally, as the researchers themselves admit, the "variations in linguistic and literary competence" among the three groups "cannot be totally effaced" (ibid.: 668).

One final study deserves discussion. After Van Peer (1986), this one is perhaps the most relevant to the current project, as it employs stimulus texts that have a Japanese connection. Belfi, Vessel and Starr (2018: 342) are interested in the "relative importance of emotional experience versus evoked imagery for aesthetic appraisals of poems" in two forms: the sonnet and the haiku. While "valence and arousal have been studied extensively as dimensions of emotional experience" in response to the visual arts, music and individual words (ibid.), the researchers wonder about the extent to which previous findings hold for poetry, and particularly for a form (haiku) which in their understanding "traditionally depends on one or two central images" (ibid.).[13] Motivating their study is the fact that "prior work has indicated that emotional valence is the strongest predictor of aesthetic appeal for visual arts [...]," but "given the purported importance of mental imagery when reading poetry," the researchers hypothesise that "poems evoking stronger vivid imagery will be considered more aesthetically appealing" (Belfi, Vessel and Starr 2018: 342).

To test their hypothesis, they secured the participation of 363 respondents using Amazon's Mechanical Turk crowdsourcing site. Respondents were split into two groups, with one group reading haiku and another reading sonnets. The set of haiku consisted of 111 texts: 68 were translated by Robert Haas from the Japanese of Bashō, Buson and Issa (these

---

[13] 'Valence' is understood as "the psychological component that evaluates [an] emotional experience" in terms of good/bad or positive/negative (Ludden 2016: 480]). 'Arousal' refers to "the physiological component of emotion that represents the degree to which the body is ready for action" [ibid]. In a sense, we may think of the two as *quality* versus *intensity* of emotion.

did not follow a 5-7-5 syllable structure in English), and 43 were haiku originally written in English by Richard Wright. (Of the Wright poems, 32 adhered to the 5-7-5 pattern, while 11 did not.) Meanwhile, the set of sonnets consisted of 16 poems ranging in date of composition from "the early 17th century to 2011" (ibid.: 343), with half in Elizabethan and half in Petrarchan form. No single participant read all the haiku—in order to "reduce participant fatigue and to ensure total reading time was equivalent between haiku and sonnet readers, each subject saw approximately one-third of the 111 haiku presented in a random order" (ibid.). Participants then rated each poem according to the following four dimensions:

1. Vividness ("How vivid is the imagery evoked from this poem?")

2. Arousal ("How relaxing or stimulating is this poem?")

3. Valence ("How positive or negative is the content of this poem?")

4. Aesthetic appeal ("How enjoyable or aesthetically appealing did you find this poem?") (ibid.).

The researchers found that across both genres, vividness was the "best predictor of aesthetic appeal […], followed by valence […], and arousal […]" (ibid.: 344). Moreover, there was "no main effect of genre on aesthetic appeal." In other words, whether a poem was a haiku or a sonnet did not seem to matter much with regard to its aesthetic appeal. These results generally run contrary to previous findings referenced in the beginning of their paper regarding visual art. Of course, as those findings had to do with a different artform, the researchers recognise that the "modality" of the art in question could be important in interpreting the results (ibid.: 345). The researchers posit three possible explanations.

The first involves the idea that "[p]oems that more readily evoke vivid imagery may help readers make sense of the poem and allow for concrete interpretations" (ibid.: 345). Remarking that a "large body of prior work has suggested processing fluency positively relates to liking," the researchers suggest that "increased vividness may improve processing

fluency" (ibid.). Processing fluency refers to "the ease with which information flows through the cognitive system (which includes both perceptual and conceptual components)" (Reber 2012: 225). This notion is implicated in the 'fluency theory of aesthetic pleasure.' As Reber (ibid.: 224) explains, the "basic idea of the fluency theory of aesthetic pleasure is simple: if people process information about an object easily, they feel positive affect, especially if ease of processing is unexpected." This explanation is intriguing in that it runs counter to Shklovsky's ([1917] 1965: 12) theory of art as energy-extensive. How might we relate the two theories? If the function of art lies in assisting people to "recover the sensation of life" through a process of refreshing their perceptions (ibid.), then one might argue that Shklovsky posited perceptual difficulty as a *sine qua non* of defamiliarisation because he experienced such difficulty in the texts he preferred to read, and never experienced defamiliarisation without it. In other words, perhaps defamiliarisation is possible *without* an appreciable negative impact on processing fluency.

Where haiku are concerned, qualities associated (at least intuitively) with processing fluency have been extolled at different times in both the Japanese context and in the reception of haiku outside of Japan. In fact, Masaoka Shiki (1867-1902)—considered one of the four greats of haiku, alongside Bashō, Buson and Issa—had even read Herbert Spencer's *The Philosophy of Style* (1852), which argued that "the more time and attention it takes to receive and understand each sentence, the less time and attention can be given to the contained idea; and the less vividly will that idea be conceived" (Spencer [1852] 1884: 11). Here, the notion of vividness is explicitly linked to "the importance of economizing the reader's or hearer's attention" (ibid.). If this sounds like the opposite of what the Russian Formalists argued, it is—Spencerian notions of style informed the two major Russian thinkers *against which* Shklovsky formulated his own theories. In any case, as Matsui Takako (1996: 28-31) relates, Shiki had esteemed Bashō's well-known frog-in-a-pond haiku as an illustration of what

Spencer called "The Principle of Economy Applied to Words" (Spencer [1852] 1884: 9), insofar as the poem succeeded in expressing the "peace and quiet of its setting" [sono-ba-no kansei-na fun'iki] without explicitly describing it in those terms (Matsui 1996: 28). Rather, it employed so-called 'minor images'; in Spencer's theory, it is not just sentence structure by which economy can be achieved, but also the "choice and arrangement of the minor images, out of which some large thought is to be built up […]" (Spencer [1852] 1884: 34). Matsui (1996: 31) suggests that the quality Shiki most valued in haiku derived in part from such passages in Spencer. The quality Shiki admired was "inshō meiryō" [clarity of impression]. As Shiki ([1897] 1979: 503) himself wrote: "'Inshō meiryō' refers to making the reader […] feel as if they can see the actual object/landscape [jitsu butsu jikkei] right before their eyes. It is thus as if [the reader] […] were looking at a small sketch." While Shiki's attitude toward Bashō was complicated and often critical, Matsui (1996: 36) argues that what Shiki ultimately took away from the frog-in-a-pond poem was the idea that "in order to craft a good haiku, one did not need to use exaggerated language [ōgesa-ni kamaeta monoii] or contrive deliberately novel [shinki-na] [scenes] […]." Rather, one could "fashion a haiku, and an excellent one at that, out of ordinary and hitherto overlooked things […]." The importance of imagery, and the sudden apprehension of an ordinary or fleeting moment seem to be two of the most frequently emphasised qualities in the international reception and practice of haiku, as well.[14] These emphases have long been implicated in the way writers and translators have understood the language of haiku.

In the Anglophone context, one of the most influential accounts of haiku was by R.H. Blyth ([1949] 1981), who emphasised "the simple directness and instantaneous perception of haiku" (ibid.: 18). Writing of ikebana, which he considered one of the "spiritual origins" of

---

[14] See e.g. the Swedish reception of haiku as described by Kodama (2008), the Argentinian reception described by Ijiri (2012), the Italian reception described by Colangelo (2012). In French, Barthes ([1978-1980] 2011: 82) wrote of the "Clarity of Haiku": "haiku=short and extremely clear: completely readable [sic]."

haiku, Blyth claimed that "[a]s in haiku, the aim is to reduce the complexity, the wild lawlessness of the material, to that point, and not beyond it, where the true nature of the thing is revealed to the poetic eye" (ibid.: 142). Here, as in Russian Formalism, the hermeneutic is one of apprehending the familiar as if it were unfamiliar—after all, the familiar perception of reality is positioned as something which literature helps us go beyond. However, the "transfer of the object to the 'sphere of new perception'" (Erlich 1980: 176) so central to the Formalist account of art is accomplished not through a "roughening" (Shklovsky [1917] 1965: 22) of language which takes longer to process, but through a kind of instant clarity supposed to derive from the subtraction of all that is unnecessary: "More must be taken away, less must be said" (Blyth [1949] 1981: 316). In this we may detect the echoes of the so-called 'instrumental' or 'conduit' model of language which post-structural thought helped to problematise (i.e. language as a transparent medium for the direct expression of thoughts). Nevertheless, so long as we are able to conceive of unexpectedness as a possibility within (or perhaps, arising from) plain, unmarked, easy-to-process language, then Shklovsky's theory would seem to have a vulnerability.

A second explanation offered by Belfi, Vessel and Starr (2018: 346) for their results is that vividness might improve attention: "readers might pay closer attention to poems that are more vivid." As they write, "[a]ttention has […] been suggested as one critical aspect of aesthetic experiences […]" (ibid.). As Pelowski et al. (2017: 108), suggest, the "level of attention taken into interaction with an artwork" seems to be related to the "activation of reward areas" in the brain. This seems partially consistent with Shklovksy's of "perception" in art and literature, particularly in Berlina's (2017: 80) more recent translation of "Art as Device": "the device of art is the '*ostranenie*' of things and the complication of the form, which increases the duration and complexity of perception, as the process of perception is its own end in art and must be prolonged."

The third explanation is that "vividness facilitates or increases emotional responses to poetry" (Belfi, Vessel and Starr 2018: 346), and that emotional responses contribute to whether readers find texts aesthetically appealing. The researchers acknowledge that their own study found "emotional valence" to be "less predictive of aesthetic appeal than vividness," but counter that they actually measured "valence of the poem *content* as opposed to *felt* emotions of the reader," qualities which previous studies show are not necessarily reflective of each other (ibid.).

Two final findings reported by the researchers are particularly relevant to my research. First, the researchers ran a regression analysis to determine "the contribution of translation to ratings of aesthetic appeal"; however, they found that while translation "did capture some of the variance, the effect of translation appears to be minimal," such that $\beta = 0.02$ (ibid.: 345). Second, the researchers found "high individual consistency and low interrater agreement" with regard to aesthetic appeal (ibid.: 347). While they admit that this could be the result of an "unreliable sample of raters"—as ever, the question of the representativeness of samples rears its head—they choose to "interpret it as an experimental result": "Participants' own judgments seem to be relatively consistent over time, whereas individuals do not agree with one another on what is aesthetically appealing when it comes to art forms" (ibid.: 345-346).

2.3. DISCUSSION: THEORETICAL AND METHODOLOGICAL IMPLICATIONS

The diversity of approaches and findings discussed above suggest the enormous complexity of literary reading and aesthetic judgment, as well as the difficulties involved in formalising relationships between textual and reader variables. However, if we arrange the studies chronologically—from, say, Van Peer (1986) to Dixon et al. (1993) to Zyngier et al. (2007) to Belfi, Vessel and Starr (2018)—we see perhaps a growing consensus that the

formalist and conventionalist approaches are actually "complementary approaches" (Belfi, Vessel and Starr 2018: 341) rather than competing ones. Moreover, in parallel with developments in cognitive science, there has been a growing awareness of what we might call the intense verticality of reading; as Wolf (2008: 10) puts it, "[a]ll human behaviours" including reading "rest on layers on layers of teeming, underlying activity," with many interrelated processes happening simultaneously or within milliseconds of each other in the brain. Insofar as there is a beginning and an end to what we read (even if these two points do not correspond to the first and last words of a text), there may always be a horizontal dimension to the act of reading. However, as I discuss later, readers do not always read in a consistently 'forward' direction, even at the motor-sensory level of eye movements. As Wolf (2008: 148) relates, "[a]t least 10 percent of the time, our eyes dart back ever so slightly in regressions to pick up past information." This iterative aspect to reading problematises the notion of a 'first' or 'second' reading, insofar as re-reading seems baked into the act of reading from the very start. (I am reminded of the comment one respondent left on one of my surveys: "Difficult to observe the instruction to read through only twice. Particularly 'What falls tonight…' where an underlying meaning is detected.") Nevertheless, this iterative aspect was frequently missing in previous accounts of literary reading such as those offered by Shklovsky or Fish (1980b), which assumed an ineluctable forward progress. While Fish ([1970] 1980: 27) argued that the category of response includes "the reversal or questioning" of "attitudes toward persons, or things, or ideas referred to" in the text, these effects are still part of "the *temporal* flow of the reading experience" [ibid.: 27], lying in wait, as it were, for the reader to experience them—and it is very much *the* reader, in a monolithic sense, that Fish is discussing.

This has implications for the design of empirical studies. If we are interested in measuring affective response as it relates to the level of perception, then we must accept the

possibility that this perception itself can be informed or altered by ongoing responses at the levels of evaluation and interpretation. The issue of discontinuities would seem especially pertinent to longer texts—it seems intuitively correct that the longer the text, the more points there would be for the reader to momentarily stop and form interpretations and judgments. However, there is nothing to stop an empirical reader from pausing in the middle of even a short poem to gather their thoughts or form an opinion about the text. Then again, the 'aha' moment of heightened awareness which Blyth ([1949] 1981) and others have emphasised as an important element of reading haiku might come only after several readings, so that we might expect an evaluation of the poem to change dramatically from one reading to the next. While I quoted a respondent above who found it difficult to read the texts only twice, it is interesting that they singled out one of the tanka translations: there are of course important differences between haiku and tanka, but the attested reaction suggests that both forms of poetry may be short enough that they encourage readers to revisit them until a kind of Gestalt is perceived. Muth and Carbon (2013: 28) refer to this phenomenon as the "Aesthetic Aha effect" in their study of whether subjects recognised faces in a series of stimulus images. The authors report that "that perceptual insights into Gestalt within difficult pictures increase appreciation" (2013: 29).

All of this points back to the difficulty of positing a single chronometry or phenomenology of literary reading. If we take the view, as Dixon et al. (1993: 14) do, that "one characteristic of literary effects is that they emerge over time," then a crucial question remains to be investigated: how much reflection is required to generate a reliable impression of literariness? After all, what may strike one as consequential in this regard upon first, second or even third reading may no longer seem so eventually, even for the most canonically literary of texts. If this is the case, then the judgment of how well a particular text belongs to some category of texts felt to be literary would seem to depend on a reader's *memory* of their

experience with the text (and of course, on their willingness to step into the role of adjudicator of literary matters, whose opinion or taste might then itself be judged). But then, what role do spontaneous impressions play? How much weight should we give to those effects that strike deepest, as it were, on a reader's first encounter? If it is a given that strong effects lessen with familiarity, should we not then promote spontaneous effects to a role of greater importance?

This last question seems particularly salient to modern reading habits. The dictates of consumerism, the compulsion to keep abreast of ever-changing cultural conversations, not to mention the exigencies of daily life—these are all surely implicated in the fact that many texts that a reader might consider 'literature' will only be encountered or engaged with once by that reader. All the while, the bedside TBR (to be read) pile grows. Thus, even re-reading (in the conventional sense) might be thought to be a historically and culturally variable practice and thus not necessarily germane to the perception of literariness in every literacy act.

If I have focused above on the difficulties of postulating qualities associated with literariness, empirical studies can nevertheless produce surprising results which confront us with the taken-for-granted-ness of our own conceptual categories. For instance, while haiku have been strongly associated with vivid imagery in a number of cultural contexts, Belfi, Vessel and Starr (2018) found that while vividness was indeed a good predictor of aesthetic appeal in haiku, it was an *even better predictor of aesthetic appeal in sonnets*. (The authors surmise that because sonnets contain more "information"—for instance, "meter, rhythm, rhyme, and syllabic and tonal variation"—this information "may aid in the production of vivid imagery" [ibid.: 346].) This points to the possibility of a discrepancy between what theorists and writers of particular genres *think* a text is doing, and what the text may actually be doing. Far from closing off avenues of inquiry, however, such findings potentially create

new perspectives from which to theorise about poetics and reception. Of course, one may argue about how representative the stimulus texts were in the above study, or question the make-up of the sample, but that is the nature of empirical research.

The foregoing may be seen as the predictable lead-up to a list of caveats concerning the present research. And certainly some caveats do need to be stated. While I am interested in responses to relatively short poems after respondents have read them twice (to make results comparable to Van Peer [1986], who issued the same instruction), the notion of re-reading is fraught to begin with. Moreover, no matter how much respondents' reading conditions resembled their preferred conditions for leisure reading (and this is highly doubtful in the case of the student respondents, most of whom completed the surveys in class), there is no escaping the artificially *prompted* nature of the survey. That said, even in an age when surveys are often administered online, paper-and-pencil surveys offer certain affordances: for one, they can reach people beyond those registered on survey websites; moreover, the paper format would seem to preserve the particular modality of reading I am interested in: the reading of physical media like books. In any case, as we look ahead to the next chapter, which covers the design of this study in more detail, this is a good opportunity to bring Translation Studies back into the discussion. As I demonstrate in the next section, many of the theoretical issues discussed in the context of Russian Formalism have parallels in Translation Studies, particularly in the work of Lawrence Venuti.

2.3.1. VENUTI AND RUSSIAN FORMALISM

Why study formalism in connection with Venuti's foreignisation? Other commentators have elaborated on foreignization's linkages and debts (acknowledged by Venuti himself) to the likes of Friedrich Schleiermacher and Antoine Berman, but fewer have explored the

similarities between foreignization and Russian Formalist theory. As I show in this section, the similarities run much deeper than Venuti's deployment of the vocabulary of estrangement (see, e.g., Venuti [1995] 2008: 177, 193, 262-264; and Venuti 1998: 85, 87, 102, 114) and defamiliarization (see, e.g., Venuti [1995] 2008: 79, 121, 161, 194, 208, 271; and Venuti 1998: 5, 12, 86, 95, 101, 115). Ultimately, these similarities rest upon broader assumptions about the nature of aesthetic response and the role of style in it, as well as the socially transformative potential of literature, and what we might call the 'variability' of readers as perceiving subjects. (I use the term 'variability' rather than unpredictability to reflect the fact that unexpected responses on the part of individual readers do not necessarily "imply a lack of systematicity and regularity at the level of process and mechanism" [Bortolussi and Dixon 2003: 23].) The similarities presented in this section not only justify the decision to revisit Russian Formalism more than a hundred years after the first meetings of the Moscow Linguistic Circle and OPOJAZ, but they also expose some paradoxical underpinnings of Venuti's theory, which in turn help us to situate Venuti's foreignisation in terms of the formalist vs. conventionalist debate described earlier in this chapter. This discussion also has important ramifications for the way I operationalise his theory in my survey research.

To state it plainly, *the way Venuti discusses foreignisation (from the 1986 article "The Translator's Invisibility" onward) resembles the way Russian Formalists discussed poetic language at different stages of that movement*. Moreover, the similarities between Russian Formalism and Venuti's foreignisation revolve around at least four concepts. That is, both poetic language and foreignisation have been conceptualised as: (1) **disjunctive**, (2) **functional**, (3) **ethical** and (4) **universalising**.[15] I discuss each of these terms in turn. I do not claim that the likeness is total—there are indeed important ways in which the theories differ. The point is that we can use the resulting similarities and dissimilarities to more fully situate

---

[15] I borrow the final term from Boase-Beier (2006: 14).

Venuti's foreignisation historically, as well as reflect on the theoretical and methodological implications of this resemblance. At the risk of oversimplifying a complex likeness, which shifts depending on which periods of the theories we compare, let me posit here one possible formulation of this resemblance: for the early Shklovsky, the verbal work of art was an autonomous "product of intentional human activity" (Steiner 1984: 50) which was rich enough to be investigated and discussed on the basis of its own intrinsic properties, and whose function was to remove "objects from the automatism of perception" (Shklovsky [1917] 1965: 13) in an ethical challenge (and therapeutic response) to those aspects of everyday life we no longer perceive, owing to the detrimental effects of habitualisation. This conception reversed the then-prevalent view that the function of poetry was to present the unfamiliar in terms of the familiar, a notion derived from Herbert Spencer's "principle of the economization of mental energy" (Steiner 1984: 49-50). Shklovsky associated this Spencerian notion with two important figures in Russian philology, Potebnya (1835-1891) and Veselovsky (1838-1906), whose theories the early Formalists would go on to challenge. More importantly, the hermeneutic movement which that notion entailed (presenting the unfamiliar in terms of the familiar), and which Shklovsky challenged, is similar to the dynamic of domestication against which Venuti's own theory is pitched. That is, Venuti seeks to overturn a similar orthodoxy that is also associated with a principle of least effort. In the context of English-language translation, that principle of least effort is "fluency as now practiced and enforced" (Venuti [1995] 2008: 19), i.e., a tendency toward "easy readability" (ibid.: 1) which threatens to reduce the unfamiliar to the familiar and render the labour of translation, not to mention translators themselves, invisible. If Shklovsky emphasised the verbal work of art as the product of an "intentional creative process" (Steiner 1984: 65) irreducible to the external circumstances of its production, Venuti emphasises the autonomy of translations as texts to be read "in their own right" (Venuti 2013: 168), whose acceptability

should no longer be judged on whether they reproduce some "univocal meaning assumed to be inherent in the foreign text" (ibid.). Thus, he endorses a translation strategy whose supposed function is akin to that of verbal art in Shklovsky's model: to renew readers' perceptions in an ethical challenge to the dominant cultural values around them. Finally, the Formalists largely assumed that readers behaved in an undifferentiated fashion. While in Venuti's theory readers assume an important role in the process of making translators' work visible, he nevertheless endorses foreignisation as a strategy that can cross over between elite and popular readerships (Venuti 1998: 12; Venuti [1995] 2008: 124) without duly considering that foreignisation itself might enact its own exclusions. There is much to unpack from this brief sketch, and I will approach it in terms of the four bases for comparison mentioned above.

First is **disjunction**. According to Peter Steiner (1984: 48), **disjunction** was the "key logical principle" by which the first Formalist model "organized its basic concepts." Advanced by Shklovsky, this Formalist model was fundamentally "mechanistic" (ibid.: 47). As Steiner (ibid.: 45) explains, "[t]echnology, that branch of knowledge pertaining to the art of human production, was the predominant metaphor applied by this model to the description and elucidation of artistic phenomena." One consequence of the machine model for Shklovsky's view of literature was that by "focusing on the nuts and bolts of poetic texts, […] mechanistic Formalism radically reversed the value of content" (ibid.: 47). If previous Russian criticism had focused on the "what" of literature (conventionally described in terms of content), then Shklovsky shifted the emphasis to the "how" (conventionally described in terms of form) (ibid.). His "inconsistent" definitions of 'form' (Erlich 1980: 187) notwithstanding, Shklovsky did not simply reverse the positions of form and content within the traditional dichotomy, but attempted to expose the separability of form and content as a "fallacy" to begin with (ibid.)—and to posit a new dichotomy in which 'device' and

'material' were the operative terms, such that both were considered elements of 'form.' Hence Shklovsky's ([1921/1925] 2017: 97) provocative formulation: "A literary work is pure form; it is not a thing, not a material, but the relation of materials."

What were these devices and materials? Shklovsky was more consistent in discussing the former than the latter. As examples of devices in poetry, Shklovsky cited the use of imagery as well as "parallelism, comparison, repetition, symmetry, hyperbole" and other such "rhetorical figures" used to increase the "impact of a thing (words and even sounds of the text itself are things, too)" (Shklovsky [1917/1919] 2017: 76-77). According to Erlich (1980: 188), "the 'materials' represented the raw stuff of literature which acquires esthetic efficacy […] and thus becomes eligible for participation in the literary work of art only through the agency of the 'device,' or, more exactly, a set of devices peculiar to imaginative literature." The 'device' was promoted to a central role in the mechanistic model, and in early Russian Formalism more generally (Steiner 1984: 50). However, the nature of the 'materials' themselves was far from clear. As Erlich notes (1980: 189), Shklovsky could endorse two competing interpretations of 'material' in the same work. Thus, for instance, in *Literature and Cinematography*, Shklovsky wavers between a conception of 'material' as external to the work of art and one in which it is nearly synonymous with language itself.

In the former conception, Shklovsky ([1923] 2008: 4) describes 'material' in terms of what a painter might take as a starting point for a painting, but which the painter is not beholden to depict in mimetic fashion. But it is the latter conception that chiefly concerns us. Shklovsky writes: "Words are not merely a means of saying something but the very material of a work of art. Literature is made out of words and comes into being by employing the laws of the word" (ibid.: 17). In a move that is crucial to our discussion of Venuti (and particularly his enthusiasm for the 'remainder,' a concept borrowed from Jean-Jacques Lecercle

[1990]),[16] Shklovsky downplays the role of the writer as the conscious force behind the organisation of a work of verbal art, arguing instead that words tend to call up other words in a kind of self-governed process.

Moreover, in Shklovsky's linking of verbal and visual art, we find an early instance of the window metaphor so central to Venuti's theory of foreignisation. Shklovsky argues that paintings cannot be stripped of form to get at their content: "Paintings […] are not windows into a different [world]—*they are things*" (Shklovsky [1923] 2008: 7).[17] Similarly, he affirms the non-transparent nature of language in poetry: "Words in poetry are not the means of expressing a thought; the words as such express themselves and they themselves, by their own essence, determine the course of a work of art" (ibid.: 9). This view of language was similar to the Russian Futurist notion of the 'self-sufficient word' (*samovitoe slovo*), and we will encounter it again below in our discussion of **function** and **ethics**. For the time being, we may note that this is "generally understood as the notion of the word as the primary fact and hero of poetry with an emphasis not on meaning but on form, texture and sound" (Cooke 1987: 68). Moreover, in its express rejection of the conduit model of language, in its alignment with experimental modes of writing, and its insistence on the autonomy of the

---

[16] Venuti (1998: 9-10) uses the term 'remainder' to refer to those linguistic items which deviate from the current "standard dialect" at any given time. Thus, he writes that "a language, at any historical moment, is a specific conjunction of a major form holding sway over minor variables" (ibid. 10). As a kind of repertoire of minor forms, then, the remainder "subverts the major form by revealing it to be socially and historically situated" (ibid.), and it serves a kind of anti-hegemonic purpose to "promote cultural innovation as well as the understanding of cultural difference." While Lecercle's (1990) conception of the remainder encompasses this function, this is by no means the only function; moreover, the remainder itself is not intrinsically politically progressive or revolutionary, as Lecercle's examples attest (see, e.g., ibid.: 194, 198, 257-258). But perhaps most crucially of all, *contra* Venuti, the remainder cannot always be marshalled consciously. As Lecercle (ibid.: 110) writes: "My conception of the remainder owes its origin to the idea that it is language that speaks [the speaker]." Finally, the remainder is connected to "the workings of the unconscious" (ibid.: 257), and as such it is paradoxically both "highly intentional" and "highly unintentional" (ibid.: 57).

[17] The translation I am citing from contains a misprint; it reads "word" instead of "world." I am grateful to Nathan Brand (personal correspondence: 1 April 2020) for confirming that the source text does in fact read "*mir*" (or "world" in English). See: Shklovsky, *Literatura i kinematograf* (Berlin: Russkoe universal'noe izdatel'stvo, 1923), p. 8.

signifier, this view of language has several parallels with Venuti's foreignisation.[18]

The disjunctive aspects of Shklovsky's model had to do with his determination to show that the history of literature was "totally immanent" (Steiner 1984: 56)—that the principles of literary change could be sought within literature itself. Against the prevailing views of his time, which sought the principles of artistic construction outside the work of art, and which subordinated the study of literature to other disciplines like politics or psychology, Shklovsky was committed to upholding a strict separation between art and *byt* (or "everyday life," as Steiner [1984: 48] translates it). This model of literary history attracted criticism from other Russian Formalists and would eventually find itself superseded by the more sociologically-minded model of Yury Tynyanov. Nevertheless, Shklovsky seemed conflicted about abandoning it entirely. As Steiner (1984: 58) notes, Shklovsky was "aware of the historical relativity of the concept of literature, but could not take full advantage of his knowledge without destroying his conceptual frame," which was based on binary oppositions. And yet, as I have tried to show above, these binaries were themselves fraught, which Shklovsky himself must have well known when making declarations like: "The most wonderful thing about [the formal method] is that it doesn't deny the ideological content of art, but considers so-called content to be a phenomenon of form" (Shklovsky [1923] 2017: 150); and the "content (soul) of the literary work equals the sum total of its stylistic devices" (Shklovsky 1921: 8, qtd. in Steiner 1984: 66).[19] As we can see from the two quotations above, despite the creation of a new disjunction between 'device' and 'material,' Shklovsky continued to use the

---

[18] Like the Russian Futurists, Venuti acknowledges that signifiers have attributes of their own which interact with each other on a different plane than that of their signifieds. The fact that this plane of interaction (between signifiers) in the TT will necessarily differ from that of the ST emphasises the ST/TT disjunction in Venuti's model. This disjunction is further emphasised at the level of function. As Koskinen (2012: 16) notes, "An important, and often overlooked, aspect of Venutian foreignization is that […] items do not necessarily need to have a direct stimulus in the source text or culture to function successfully in their foreignizing task (see Venuti 1995: 290–292)."

[19] I am grateful again to Nathan Brand (personal correspondence: 2 April 2020) for confirming that the Russian word for 'content' in both of these quotations is the same: *soderzhianiye*. The source for the first instance is Shklovsky, *Sentimental'noye puteshestviye* (Berlin: Gelikon, 1923), p. 327. The source for the second instance is Shklovsky, *Rozanov* (Petersburg: OPOJAZ, 1921), p. 8.

terms 'content' and 'form,' and moreover defined them in paradoxical ways: in one formulation, form subsumes content, while in another, content subsumes form. This unity-within-dichotomy extended to his new terms 'device' and 'material.' While Czarniawska (1997: 47) applies the terms in a different context, I believe she sums up the problem well when she writes that "[d]iscussing any material presumes a device, otherwise no discussion can take place, as there is no such thing as formless material."

A similar air of paradox attends Venuti's domestication and foreignisation. In a theoretical move reminiscent of Shklovsky, Venuti opposes domestication and foreignisation to previously existing dichotomies in translation discourse (including "'literal vs. free,' 'formal' vs. 'dynamic,' and 'semantic' vs. 'communicative'") while insisting that his new terms do not in fact "establish a neat binary" at all (Venuti [1995] 2008: 19). As he tells us in the preface to the third edition of *The Translator's Invisibility* ([1995] 2018: xiii), "to treat the distinction between domesticating and foreignizing translation as a simple 'dichotomy' or 'binary opposition' is to eliminate entirely its conceptual complexity." There is complexity to be found in his theory, to be sure, particularly when he writes that the "significance of translation practices is always historically contingent," and that individual translations can be identified as "'foreignizing' or 'domesticating' […] only in relation to specific cultural situations, specific moments in the changing reception of foreign literatures and in the changing hierarchy of values in the receiving culture" (Venuti [1995] 2008: 232). But complexities and inconsistencies are two different things, and Venuti's treatment of his main concepts results in some "logical difficulties," as Tymoczko (2000: 34) calls them. Just as for Shklovsky there could be no form without material, for Venuti there can be no foreignisation outside of domestication. Why should this be the case? For Venuti, translation "always involves a process of domestication, an exchange of foreign-language intelligibilities for those of the translating language" (Venuti [1995] 2008: 177). Or as he puts it more forcefully

in the preface to the third edition of that work: "**All translation, regardless of genre or text type, including translation that seeks to register linguistic and cultural differences, is an interpretation that fundamentally domesticates the source text**" (Venuti [1995] 2018: xii). Thus, we are forced to conclude that foreignising translation is a subset of domesticating translation; moreover, the set of texts called 'domesticating translations' is co-extensive with the set consisting of all translations. This in itself is perhaps not so problematic. The real difficulty appears when we realise that he is using the term 'domesticating' as both a superordinate term and a subordinate term, as the figure below shows:

Fig. 2. Domesticating and Foreignising Translation
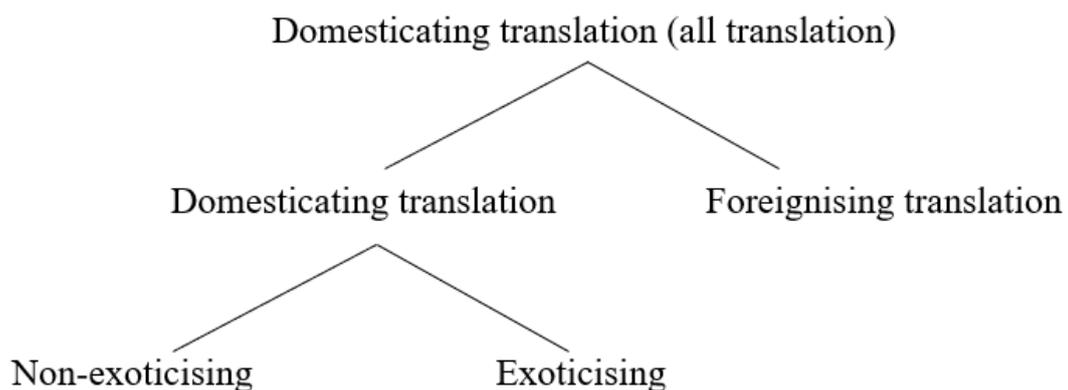
Whereas the superordinate term 'domesticating' refers to a fundamental and inescapable aspect of translation in general (according to Venuti's model), the subordinate terms 'domesticating' and 'foreignising' seem to refer to a different order of phenomenon. Indeed, they seem to refer to the "overall impact" of a translation (Venuti [1995] 2018: xiii). Regarding the issue of dichotomy, Venuti ([1995] 2018: xiii) writes: "I would want to distinguish, at the very least, between translations that are domesticating, exoticizing, or foreignizing in their overall impact, where 'exoticizing' registers a superficial sense of difference that can easily play into cultural or ethnic stereotypes." In terms of schematisation,

I will not linger too long on 'exoticizing'; the way Venuti ([1995] 2008: 160-163) discusses it elsewhere seems to mark it out as a subset of domesticating translation in the subordinate, 'overall impact' sense of the term. (I have added a 'non-exoticising' branch to the diagram above for the sake of completeness: if some domesticating translations exoticise, then it follows that others do not.) Where does this leave us? Still stuck with the two main terms: domesticating and foreignising. Thus, one of the major problems that remains to be solved in Venuti's theory is the relationship between the first and second tier of the above diagram, that is, between domestication as a superordinate term and domestication/foreignisation as subordinate terms. This problem is compounded by the fact that Venuti does not just differentiate between domestication and foreignisation in terms of overall effect—he has also defined foreignisation in a *negative relation* to the superordinate sense of domestication. In this formulation, a translation foreignises to the extent that it "limit[s] and redirect[s] its inevitable domestication" (Venuti [1995] 2018: xiii). This problem of degree suggests that it is appropriate to think of foreignisation as a gradient phenomenon. But if we put foreignisation and domestication (in the subordinate, 'overall effect' sense of the terms) on either end of a spectrum, what would it mean for a text to fall exactly midway? Would this mean that it is *neither* domesticating *nor* foreignising in overall impact—or perhaps both? Or would we have to come down on the side of domestication, given that all translation for Venuti (ibid.: xii) is "inevitably domesticating" (in the superordinate sense of the term)? There does not seem to be a way out of this impasse without a redefinition of terms on Venuti's part, or at least a clarification about how these notions interrelate.

In terms of the formalist vs. conventionalist debate, Venuti's emphasis on "overall impact" (ibid.) as the relevant criterion for the application of these terms suggests a reader-reception model, in which we should examine the responses of readers, either individually or in the aggregate, for evidence of the effects of foreignisation. That is, the ultimate arbiter of

such effects would be readers, aligning Venuti's theory with the conventionalist model of reading. I would argue this is very much the mode in which Venuti discusses translation in the 1986 essay "The Translator's Invisibility." And I would argue this is true even if the method of reading he endorses as a corrective to the translator's invisibility severely circumscribes the potential to affect "existing class relations" (see Venuti 1986: 189-190). In short, Venuti (1986: 197) is concerned with how the translator's work of "transforming the original" can be perceived by readers who are unable to compare the ST and the TT because they do not have the "advanced proficiency in several languages necessary to pursue such a rigorous line of research." He asks: "Is there some way for sensitive, informed readers to glimpse the transformative process?" (ibid.). The answer is yes, and in order for this to be accomplished, he endorses what he will later call "symptomatic reading" (Venuti 1995: 25). The method of reading called for here is "one that tries to be more thoughtfully responsive to those noticeable discrepancies that have hitherto been regarded simply as defects—to logical flaws in the choice of words, for example" (Venuti 1986: 198). By the first edition of *The Translator's Invisibility*, he has elaborated on what a symptomatic reading entails: such a reading "locates discontinuities at the level of diction, syntax, or discourse that reveal the translation to be a violent rewriting of the foreign text, a strategic intervention into the target-language culture, at once dependent on and abusive of domestic values" (Venuti 1995: 25).

Performing a symptomatic reading on the English translations used for Freud's *Standard Edition*, Venuti (ibid.: 29) claims that such a reading "can be said to foreignize a domesticating translation by showing where it is discontinuous." (This language is also present in the second and third editions of *The Translator's Invisibility*.) While this claim would support the argument that Venuti's theory is conventionalist, insofar as it portrays foreignising effects to be principally reader-directed, it occurs in a body of work which emphasises time and again that it is chiefly the intervention of the translator that is

responsible for those effects. Sometimes these claims for the decisive role of the translator are themselves internally inconsistent. For instance, in *The Scandals of Translation* (Venuti 1998: 68), he seems to want it both ways:

> Whether the effects of a translation prove to be conservative or transgressive depends fundamentally on the discursive strategies developed by the translator, but also on the various factors in their reception, including the page design and cover art of the printed book, the advertising copy, the opinions of reviewers, and the uses made of the translation in cultural and social institutions, how it is read and taught.

What makes this passage difficult to interpret is the word 'fundamentally.' What is its scope as a modifier? Does it span across the conjunction 'but'? The list of reception factors is so long, and their types so various, that it is difficult to see how the responsibility for effects *could* reside 'fundamentally' with the translator or their 'discursive strategies.' Nevertheless, the claim for the determining role of the translator is echoed in another passage: "A translator can signal the foreignness of the foreign text, not only by using a discursive strategy that deviates from prevailing discourses […] but also by choosing to translate a text that challenges the contemporary canon of foreign literature in the translating language" (Venuti [1995] 2008: 126). The fact is, the criteria for describing a translation as foreignising in Venuti's theory frequently shift—an issue to which an I return in section 3.1.

   If we recall from earlier in this chapter the conventionalist preoccupation with proper interpretations (e.g. Culler's [1975] notion of literary competence), then in symptomatic reading we have an example par excellence. Note how Venuti (1986: 197) already qualifies the kind of readers who can perform symptomatic reading as "sensitive, informed." It would

not be too much to say that the theoretical background required to perform a symptomatic reading *à la* Venuti would be found primarily in a cultural elite. Yet if it is in the interest of such an elite (and particularly those involved in publishing) to keep the work of translators invisible, owing to the "consumability" of fluent translations on the book market (see Venuti 1986: 187), why would we expect anyone outside of academia to adopt symptomatic reading, especially if the aesthetic expectations of such readerships were already being satisfied? Recently, Venuti ([1995] 2018: xvii) has implicitly contrasted symptomatic reading and reading for enjoyment. As he writes, even those readers "who know the [source] language are likely to be reluctant to pursue" a symptomatic reading because "their interest in the translation may be limited to readerly pleasure" (ibid.). And yet, he argues, translation as a distinct text type demands this extra effort: translations should not "pass" for their source texts, they are different and "should be read differently, even if they require the development of a new kind of literacy" (ibid.).

This new "literacy" of Venuti's (ibid.) resembles Culler's (1975) notion of "literary competence" not just in terms of a conventionalist emphasis on expertise, but in a way which returns us to the realm of Russian Formalism and the principle of disjunction. Both Venuti and Culler insist on a disjunction concerning text type. In Culler's case, the disjunction resembles one important to the Russian Formalists: the difference between literary and non-literary language. However, as befits his conventionalist stance, the locus of this difference resides outside the text for Culler (see e.g. Culler 1975: 114). Similarly, Venuti's call for translations to be read as translations also implies a disjunction between text types: the disjunction between any given ST and TT, as well as the disjunction between non-translated and translated literature in general. However, the problem here is not *whether* readers are familiar with some set of pre-existing conventions for reading translations. Rather, the very perceptibility of translations as texts to be read "in their own right" (Venuti 2013: 168)

depends on readers *not* following existing conventions. In other words, this perceptibility

depends in part on readers learning to overcome or reject their previous reading conventions

to adopt others, in a kind of readerly self-denial (or at the very least delayed gratification).

While Venuti ([1995] 2018: xvii) argues that the pursuit of "readerly pleasure" is not a

sufficient way of engaging with translations, and that "both elite and popular readers must

learn how to read a translation […] as an interpretation" (Venuti [1995] 2008: 124), it is

through an appeal to another level of appreciation that he endorses this new kind of literacy:

> Readers can increase their appreciation of translations by deciding not to read
> them as isolated texts. They can rather create their own contexts of
> interpretation by joining their experience of a particular translation with other
> translations from the same or different source languages as well as with
> original compositions written in the translating language. Such contextual
> reading can help to make the translator's interpretation visible, provided that
> readers broaden their focus to include patterns in the selection of source texts
> while attending to the textual features of the translation itself, its cultivation of
> dialects, styles, and discourses that are rooted in the translating language and
> culture.

(Venuti [1995] 2018: xviii)

Thus, it seems we are in the midst of another disjunction: this time between two different

kinds of appreciation. The downplaying of "readerly pleasure" in Venuti (ibid.) in favour of a

new specialist mode of reading recalls a passage of Bourdieu's "Outline of a Sociological

Theory of Art Perception," in which the sociologist posits two "forms of aesthetic pleasure"

at opposite ends of a spectrum ([1968] 1993: 220). At one end of the spectrum, there is "the

enjoyment which accompanies aesthetic perception reduced to simple *aisthesis*"; 'aisthesis'

here refers to response at the "level of sensations and affections" (ibid.). And at the other end

of the spectrum, we have "the *delight* procured by scholarly savouring, presupposing, as a necessary but insufficient condition, adequate deciphering" (ibid.). Bourdieu refers to the former mode rather grimly as an "inferior and mutilated form of the aesthetic experience" (ibid.: 219), as it cannot relate the individual work "to the ensemble of the works forming the class to which it belongs" (ibid.: 222). Meanwhile, 'delight' depends on something that resembles Culler's literary competence, a special background knowledge which in Bourdieu's model is the "result of a long process of inculcation which begins (or not) in the family, often in conformity with its level of economic, academic and cultural capital" (Johnson 1993: 23).

From the standpoint of survey operationalisation, it is fortunate that these two types of response can be broadly characterised as affective and cognitive, respectively. In advocating a new literacy which affirms translation as its own text-type, Venuti seems to acknowledge that foreignisation has a potential cost in terms of affective response, which suggests we should look out for a particular pattern in our data—*a lower overall rating on affective measures as a potential parameter of foreignisation*. Moreover, we can already make the following observation: in describing effects on readers, Venuti's treatment of translated literature draws upon a much older philosophical-aesthetic tradition in which the enduring value of literature derives in part from the way it engages not just the emotions but the higher faculties, as well. We may think here of Horace's dictum that poetry should instruct and delight, an idea whose influence reached far beyond its own time to the Renaissance and beyond.[20] While the degree to which the so-called 'Horatian platitude' is still alive and well in an explicit form in contemporary pedagogy may be a matter of debate, it is probably not too controversial to say that in the Anglophone world the idea that literature is something of lasting value or importance, as opposed to merely fleeting entertainment, *has been* influential

---

[20] For a discussion of its role in Renaissance and neoclassical thought, see, e.g., Cronk (1999: 199-204).

and even common-sensical. The fact that so many English dictionary definitions of 'literature' still include descriptions like the above testifies to the staying power of this idea, even if only as a trace. Somewhat more recently than Horace, the philosopher R.A. Sharpe [1983: 95] has written that "[w]e assume that the pleasures of art ought to last," and that a "great novel will repay re-readings […]." While the 'we' in the above could use some specification, it is a common enough opinion that great art appeals to more than the baser instincts or emotions and rewards close scrutiny of the kind Bourdieu gets at in his term "delight" ([1968] 1993: 220). Whether or not this characterises the behaviour of actual readers is another question; nevertheless, in this view the temporal arc of appreciation tends toward the cognitive.

What is interesting for my purposes is that the relationship between cognitive and affective components in Venuti's framework of response to translated literature is basically similar to the framework of response to *non*-translated literature in the Horatian tradition. If we can use the term 'literary' here to describe a reading experience arising from the particular configuration of affective and cognitive responses discussed above, then we might say that at this level of analysis translated literariness and literariness do not seem all that different. But then, insofar as this view privileges the cognitive, it is different from that of Shklovksy, who treats literature as "primarily engaged with sensations" (Askin 2017: 172). So then: where does this leave us in terms of situating Venuti in the formalism vs. conventionalism debate? Though Venuti's foreignisation and Shklovsky's mechanistic model share a similar principle of disjunction and some of the same problems arising from that principle, Venuti's description of the determining role of readers seems to align his theory more closely with the conventionalist position of literary processing. However, this will change when we look at **function** and **ethics**.

In one of Shklovsky's most well-known formulations (in the seminal essay "Iskusstvo

kak priem," or "Art as Technique" in Lemon and Reis's translation), the **function** of art is to "remove objects from the automatism of perception" (Shklovsky [1917] 1965: 13). But it is not just our perception of the world that art refreshes. It also renews our feeling for language itself, as he argues in an earlier essay, "Resurrecting the Word" (Shklovsky [1914] 2017: 64). If it was the case for Shklovsky that "[t]oday, words are dead, and language resembles a graveyard" (ibid.: 63), he credited the work of resurrection to Futurists like Khlebnikov, Mayakovsky and Kruchenykh, whose neologisms and experimentation answered the need for the "creation of a new […] language […] aimed at seeing instead of recognizing" (ibid.: 71-72). As in Venuti's theory, visibility is used as a kind of metonymy for the meaningful presence or perceptibility of something previously taken for granted. But it is notable that even in contexts where a 'hearing' word might be expected, Shklovsky uses a 'seeing' word. Thus, for instance, Shklovsky writes in the same essay: "When, in a fit of affection or rage, we want to say something tender or insulting, then we are not content with worn-out, gnawed words, and so we crumple and break words to make them touch the ear, to make them seen and not recognized" (Shklovsky [1914] 2017: 70). Berlina (2017: 70n12) explains that in the final clause Shklovsky indeed uses "the word 'seen' (*uvidali*), not 'heard,' though he is talking about oral speech […]."

In contrast to Potebnya, for the early Shklovsky, the important thing about poetry was not what its imagery might tell us about the writer's life and times; rather, it was that poetry transforms the way we see the world today (the here-and-now).[21] In other words, poetry is less about (figuratively) looking inward than (literally) looking outward. In this way, Shklovsky posited for poetry—and literature in general—an agentive role in social life based on a conception of 'seeing' as an "active, dynamic act of perception brought into play by the

---

[21] Of course, as Steiner (1984: 55) notes, this implied a "relationship between literature and everyday life" which Shklovsky's mechanistic model could not support. In other words, though he strove to maintain a strict separation between art and *byt*, Shklovsky's functional notion of poetry implied that the "value of art [was] a function of its utility for *byt* […]."

artist's technique which allows us to see what, until then, had not and could not yet come into view," as Benjamin Sher (1990: xv) writes. If we set aside Sher's (1990: xvii) questionable assertion that Shklovsky saw the artist as some sort of "man-god," and focus instead on the utopian rejection of "the old" in favour of "the new" (ibid.) in Shklovsky's theory, then we begin to see how crucial the interplay of diachrony and difficulty are to the notion of an ethical poetry. For it is clear that for Shklovsky, the renewal of perception attained through the "roughening of form" (Fry 2012: 84) in poetry has **ethical** consequences. As O'Toole (2001: 165) writes: "[…] [defamiliarization] was not just a technical matter for Shklovsky; it becomes an ethical issue. One reason why we should take literature very seriously indeed is that it refreshes and renews our vision of reality; it resurrects not just 'the word,' but *the world*." He goes on to write: "Ethically, this deliberate 'putting the brakes on' perception is part of the 'resurrection of the word' and the revival of moral awareness" (ibid.: 166). While this aspect of Formalist theory tended to fade into the background as the Formalist programme became more linguistically sophisticated, the dimensions of this moral awareness were immense in the early period. The effects of habitualisation (or "automatization" in Berlina's translation) could indeed be a matter of life and death. In "Art as Device," Shklovsky quotes an entry from Tolstoy's diary in which the novelist observes that a life lived in thrall to unconscious routine has not really been lived. Shklovsky remarks afterwards: "This is how life becomes nothing and disappears. Automatization eats things, clothes, furniture, your wife and the fear of war" (Shklovsky [1917/1919] 2017: 80).

However, there is a sense in which translation may impose additional space between readers as perceiving subjects and readers as social actors, making the above ethical function a merely possible one rather than a guaranteed one. This seems to be related to a felt lack of expertise when confronted with a text which comes from another linguistic and cultural milieu—though, to be sure, the spectre of expertise may also haunt evaluations of texts which

readers think are originally TL texts. (For instance, one respondent to Survey 3—the survey in which the texts *were not presented as translations*—remarked that they found it difficult to answer the questions involving 'literariness' through their own "ignorance.") However, the difficulties attested by respondents seemed more pronounced when the texts *were* presented as translations (judging by the degree of elaboration of the comments). For example, one respondent wrote: "I found two of the questions really difficult: *Is this poem an example of good literature?* and *How literary is this poem?* It depends on what you think qualifies as 'good literature' or 'literary' and if indeed you feel well-read/qualified enough to have an opinion. I know what I enjoyed reading, and why I enjoyed it, but felt very uncomfortable being asked to judge its 'literariness.'" Another respondent put it in starker terms: "I don't know what 'good', 'literature' or 'literary' means. I don't know how any poem relates to the words 'good', 'literary', and 'literature.' […] I suggest that you attempt to access whatever it is you are after through a less direct approach. Or find an approach that doesn't assume mutual definitions of ambiguous/contested words?" Of course, I had not assumed mutual definitions, but had taken a similar tack to the Royal Society of Literature's (2017a: 9) "Literature in Britain Today" survey: "It was acknowledged from the outset that 'literature' (as distinct from other kinds of written work) is a term intrinsically open to varied interpretation – indeed, that one of the questions for the research was to see what collective meanings might emerge – and so it was decided to empower the survey respondents to use whatever definition of literature they chose." Given that I had emphasised in the survey instructions that "[t]here are no right or wrong answers" and that I was "interested in how *you* have responded to each poem," it is all the more tempting to read such responses as symptomatic of a split between what Toth (2019) has called the "reading self" and the "self-in-the-world."

In Toth's (ibid.: 11) model, readers are "made up of a reading self and a self-in-the-

world." Building on the idea of reading as both "intrinsically textual and material" (ibid.),

Toth argues that reading "splits the subject, manifesting a textually-oriented reading self and

a materially-oriented self-in-the-world" which "exist in mutually-transformative dialectical

tension" (ibid.: 18). On the one hand, the reading self is "our imaginative actualisation of the

roles offered up to us by the text" (ibid.: 53), and it "accounts for the textual dimensions of

reading" (ibid.: 54). Meanwhile, the self-in-the-world "denotes the reader as a […] corporeal

agent" who has "material and epistemic coordinates as well as memories of experience"

(ibid.: 52-53). As Toth (ibid.: 55) explains: "Together, the reading self and the self-in-the-

world occupy and move between the material and the textual. We might visualise this as the

text requiring that the reader maintain one foot in the world whilst stretching the other foot

out to the fictional or unknown universe." The mechanism by which this happens is laid out

in the following terms: "The text affects the reading self. The reading self affects the self-in-

the-world. The self-in-the-world affects the reading self. The reading self affects the text.

This happens again and again over the course of reading" (ibid.: 15).

Earlier, in section 2.1.1, I suggested that in the absence of an authority to (pre)validate

their responses, readers like the ones cited above seemed to want to preserve a buffer between

the self that experienced the text and the self that was commenting upon it. These two selves

seem to fall more or less into the two categories posited by Toth (2019): the reading self and

self-in-the-world. I also wrote that such responses suggested a kind of plausible deniability of

literary response: it seemed that, in leaving a record of their evaluations of a given text

(however anonymously), some respondents may have felt compromised in a curious way—

aware of the possibility of having their responses de-legitimised or contradicted after the fact

by someone with more expertise in Japanese poetry and/or the idea of 'literature'. In this

sense, the discomfort these respondents felt may be read as a demurral, an expression of the

tentative character or even reversibility of their responses, and an affirmation of the

possibility of a *more appropriate response* (to be instantiated by a re-reading at some future point), attendant upon further knowledge of the source poetry. Indeed, after completing the surveys, several respondents indicated a desire to learn more about the poets featured in the surveys, or about poetry in general.

If we may characterise affective engagement with a text as ethical (inasmuch as we respect the other enough to willingly yield a portion of control over our affective state), and if we may characterise an awareness of the situated-ness of reading as political, then the responses cited earlier suggest both an ethical and political engagement with the stimulus texts. In Toth's (2019) model, it is the reading self that is affected by the text. Moreover, the reading self "transmits affects, which make possible ethical interfaces like empathy, which may bring about imperatives for self-reconstitution" (ibid.: 55). Why the focus on 'self-reconstitution'? In Toth's (ibid.: 50) account, the "guiding principle" of reader response theory is that "reading entails […] experiencing the other as the self, and the self as the other in the co-construction of meaning." Toth (ibid.) is not alone in characterising this operation as a matter of ethics. As Culler (2011: 123) writes: "If one of the oldest ethical injunctions is 'Do unto others as you would have them do unto you,' a modern version, in what is called the 'ethics of alterity,' is to respect the otherness of the other." In other words, "[d]o not presume that the other is like you and thus claim a posture of universality, as if your views and feelings had universal validity" (ibid.). Reading is thus "ethical" inasmuch as it "provides opportunities for (self-)transformation and (self-)reconstitution" (Toth 2019: 304). However, Toth (ibid.) continues, reading is also "political": "It encourages readers to recognise themselves, and to be sensitive to the material and epistemic circumstances that condition their reading – from the privilege of literacy and English-language fluency, to their cultural and institutional associations." While the "ethical and political opportunities" offered by reading are "not always taken up" (ibid.), the awareness of *potentially more appropriate*

responses signalled by the above respondents can be read as a form of both ethical and political engagement. If ethics "requires an openness to other possibilities" (Culler 2011: 122), then the openness to re-reading at some future date is quintessentially ethical. However, insofar as it is premised on the validating authority of expertise—an expertise that will not be available to all readers—it would seem to "optimistically celebrate the possibilities of highly professionalised readings" (Toth 2019: 14). Meanwhile, we have seen how some of the readers cited above have resisted ordering their experience in terms of the linguistic categories used in my Likert questions—and have indeed sought to alert me to the contested-ness of terms like 'literature' or 'literary.' This is the action of readers sensitive to their own "cultural and institutional associations" (ibid.: 304), and as such it is inherently political.

But how does ethical and political engagement at this level transform into the kind of macro-level change discussed by both Shklovsky and Venuti? Particularly on the part of those readers "who tend to be denied access to literary culture's value regimes" (Toth 2019: 9)? While Venuti ([1995] 2008: 277) admits to a "utopian faith in the power of translation to make a difference, not only at home, in the emergence of new cultural forms, but also abroad, in the emergence of new cultural relations," it is difficult not to hear this as a cross-cultural elaboration of Shklovsky's own grand pronouncements on the role of literature in social life. But because their theories do not address the interface between the micro and the macro, there is a frequent equivocation between means and ends.

As far as Venuti is concerned, the attainment of visibility—of the translator, of TTs as "texts in their own right" (Venuti [1995] 2008: 13)—is good in and of itself (i.e., it is an end), but it is also a way of effecting change (i.e., it is a means to an end). As an end in itself, the visibility of the TT becomes part of the very definition of "[g]ood translation" for Venuti (1998: 11):

> Good translation is demystifying; it manifests in its own language the

foreignness of the foreign text […] This manifestation can occur through the selection of a text whose form and theme deviate from domestic literary canons. But its most decisive occurrence depends on introducing variations that alienate the domestic language, and, since they are domestic, reveal the translation to be in fact a translation, distinct from the text it replaces.

But visibility is also a means to an end: "The goal is ultimately to alter reading patterns, compelling a not unpleasurable recognition of translation among constituencies who, while possessing different cultural values, nonetheless share a long-standing unwillingness to recognize it" (ibid.: 13). In principle, there is no reason why visibility cannot be pursued both as an end in itself and as a means to further social change. Indeed, to the extent that Venuti's theory may be "construed as a moral injunction to 'stand up' for one's rights" (Atwell 1986: 105) as a translator, it is consonant with Kant's so-called 'principle of humanity': "Act in such a way that you treat humanity, whether in your own person or in the person of any other, never merely as a means, but always at the same time as an end" (qtd. in Atwell 1986: 105). However, without a better understanding of *how* individual acts of translation and/or reading lead to social action—social action which would, for instance, improve the conditions of translators or promote more "democratic geopolitical relations" (Venuti [1995] 2008: 16)— the grounds for expecting that these acts *will* lead to social change are not well established.

Means and ends are furthermore accompanied by intentions in a kind of three-way tug-of-war in Venuti's model, particularly when it comes to providing the basis for describing translations as fluent or resistant, domesticating or foreignising. As he writes in *Scandals of Translation*, the "key issue is not simply a discursive strategy (fluent or resistant), but always its intention and effect as well" (Venuti 1998: 188). This leads to a conundrum similar to the one we saw in early Russian Formalism, pertaining to the ontological status of texts which no longer perform the function which once defined them as literature. As Steiner (1984: 56)

observes, "[w]ithin the framework of mechanistic Formalism this category is a conceptual bastard, in that it is composed of artworks whose form, paradoxically, is not perceptible." Where Venuti is concerned, we might ask: should a translation that is foreignising in intention but not in effect really be called foreignising?

Implicated in this problem is the above-mentioned 'ethics of alterity.' If we conceive of this ethics in brutally simplified terms as a kind of respect for otherness, then broadly speaking, there are two ethical philosophies that one may use as a basis for describing a translation as ethical with regard to alterity: 1) intentionalist (broadly speaking, this ethical school holds that "it's the thought that counts" (Graham 2011: 89), and 2) consequentialist (it is not the thought that counts, but the result—in other words, 'the proof is in the pudding'). It is not an exaggeration to say that Venuti is basically concerned with the social consequences of foreignisation (see e.g. Venuti [1995] 2008: 13). However, if it is consequences he is after, then it seems somewhat counterproductive to reserve the right to deem a translation foreignising on the basis of intention.

The ethics of alterity might seem more likely to come up in a theory of translated literature than a theory of non-translated, 'original' literature. But an ethics of alterity underwrites the very mechanism of *ostranenie* in Russian Formalism, and marks another point of comparison with Venuti's theory. In proposing that 'poetic language' revises readers' perspectives of the world and themselves, the Russian Formalists acknowledged the potentially transformative impact of the language use of the other on the self—even if the self and other are users of the 'same' language.[22] The Russian Formalists thus recognised that the kernel of alterity that made estrangement possible lay in language itself—in the fact that there is always another word or phrase that one might have chosen. While they never agreed upon a model for the cumulative effect of these choices, the sense of an encounter with an other

---

[22] I am referring to 'language' here as a kind of shared communal resource. Cf. Chomsky (1988: 36).

117

looms large: subsequent scholars and researchers have emphasised the role of literature as a kind of staging ground for alternative relationships between the self and others (see e.g. Miall 2006: 17 and Hogan 2004: 116). To the extent that we are involved or invested in a particular text, literature may give us, however temporarily, the sensation of taking on other peoples' consciousnesses, of inhabiting other identities.

In this sense, the Russian Formalists were alive to the possibility of otherness *within* linguistic communities; indeed, the idea of estrangement depends upon it. However, if interlingual literary translation brings with it the sense of additional ethical obligations (embedded as it often is in networks of asymmetrical cultural exchange), this may be also because it implies *degrees* of otherness. Whereas the Russian Formalist models affirm that anyone who is not oneself is other, the notion of interlingual translation (particularly when understood as a kind of spatial transfer) implies that some others are *more other* than others. Where this latter otherness is concerned, Venuti (1998: 11) proposes to bring readers to an ethical awareness of the "autonomous existence" of the ST through exposure to non-standard linguistic variations in the TT, in a kind of translation practice he calls "minoritizing" (ibid.). As Venuti explains: "Any language use is [...] a site of power relationships because a language, at any historical moment, is a specific conjuncture of a major form holding sway over minor variables. Lecercle (1990) calls them the 'remainder'" (ibid.: 10). Venuti continues: "Good translation is minoritizing: it releases the remainder by cultivating a heterogeneous discourse, opening up the standard dialect and literary canons to what is foreign to themselves, to the substandard and the marginal" (ibid.: 11).

I mentioned at the beginning of this section that like Shklovsky, Venuti seeks to overturn an orthodoxy associated with a principle of least effort. In Shklovsky's case, the offending notion was ultimately Spencer's "principle of the economization of mental energy" (Steiner 1984: 49-50). For Shklovsky the opposite was true: the "perception of art manifests not the

law of least effort but the law of maximal effort," as Steiner (1984: 50) writes. Thus, Shklovsky's notion of the device entailed presenting the familiar in terms of the unfamiliar. Venuti's theory advocates a similar hermeneutic movement, though it paradoxically places the foreign text in the 'familiar' category, and the minor domestic linguistic varieties in the 'unfamiliar' category. The justification for this is not well elaborated, although Venuti (1998: 84) affirms that "domestic canons for foreign literatures are always already in place when a translation project is developed." This point could probably be disputed when it comes to languages infrequently translated into English: is there really a 'canon' of Faroese literature in English? Moreover, if there were, how many English readers could claim to be familiar with it—to the extent that their sense of this literature would need to be re-formed, as opposed to formed in the first place?

This is one aspect of the **universalising** tendency of Venuti's theory. For the hermeneutic motion described above to work, at least at the level of linguistic strategies, Venuti's theory would seem to assume a certain blanket degree of familiarity with the foreign literature on the part of readers. Otherwise, the theory would entail presenting the unfamiliar in terms of the unfamiliar. This has consequences for the ethics of representation of foreign cultures in receiving scenarios; specifically, it would seem to forestall representation of the foreign culture and privilege the representation of domestic diversity. Robinson (2017: 141-2) employs the medical terminology of "titrations" to explore the consequences of different admixtures of familiar and unfamiliar. As he argues, "we trust translators not only to *represent* [foreign] cultures to us by creating cribs to them, but to *realize* those cultures, to make them feel real to us, by recreating foreign texts in ways that we experience as a new affective-becoming-conative channel or conduit of reality" (ibid.: 139). This must be understood in the context of his notion of "icosis": this is "the *communal shaping* of individual understanding" (Robinson 2013: 12) which proceeds by a process of "group

plausibilization" (Robinson 2017: 39). As he explains, the "idea is that reality, truth, identity, the self, and so on […] are 'plausibilized,' normativized as real, by the circulation through the somatic exchange of value," i.e. through people in groups sharing affects and affective evaluations (ibid.). Nevertheless, Robinson (2017: 153) asks, "what happens when the Feeling of the Familiar is suddenly vaporized out of the dosage?" In Robinson's model, an overdose of the unfamiliar may have unpleasant side effects: "the resulting depletion of felt connection with individualized collective reality can flip us over into disturbing or disorienting depersonalization" (Robinson 2008: 99-100).

Where the representation of foreign cultures is concerned, this kind of over-*un*familiarity may foreclose the possibility of identification with the other altogether. Inasmuch as Kant's principle of humanity is implicated in an ethics of alterity, Venuti's emphasis on a translation practice that "preempts the reader's identification" (Venuti 1998: 12) so that they do not confuse the TT for the ST seems to miss the mark. It is difficult to imagine an ethics of alterity that *does not begin* with some identification of the self with the other, à la Kant. Some identification of the self with the other would seem to be a *sine qua non* of an ethics of alterity. It could be argued that this is the foreignising translator's very starting point—it is their adherence to the humanity principle that motivates the search for strategies which defend the difference of the other, and which seek to prevent this difference from being mobilised merely as a means (i.e. commodified) rather than as an end in itself. However, in arguing against readers' "sympathetic identification" (Venuti [1995] 2008: 163) with cultural others, Venuti assumes that "sympathetic identification" and "critical detachment" (ibid.) are mutually exclusive modes of reading, and furthermore, that popular readerships are particularly incapable of the latter. As he writes: "Popular literary genres might be considered as particularly susceptible to exoticizing effects because they work to solicit the reader's sympathetic identification so as to produce the realist illusion, an unreflective response that

lacks critical detachment" (ibid.: 163).

In this way, Venuti's stance resembles that of the materialist postcolonial scholars criticised by Toth (2019: 20) for their "demonisation of non-academic readers." She cites Brouillette's (2007: 25) observation that "more elite readers often seem to believe that there is a general public engaged in a similar activity (reading), but who practise it *badly*" (qtd. in Toth 2019: 19). Moreover, the supposed susceptibility of popular audiences to fall victim to the illusion of unmediated contact with the other *is a trait Venuti later singles out with respect to elite readers* ("theorists, critics, and translators") in his discussion of "simpatico" (Venuti [1995] 2008: 237-238). Therein, Venuti recounts the advice he once received from "an older, widely published, and very gifted writer": "The translator works better when he and the author are *simpatico*, said my friend, and by this he meant not just 'agreeable,' or 'congenial,' meanings which this Italian word is often used to signify, but also 'possessing an underlying sympathy.' Venuti (ibid.: 238) goes on to say that "my friend's ideas about translation still prevail today in British and American cultures […]."

Not only does Venuti distinguish popular readerships from more specialised ones *on the basis of a trait also present in specialised readerships*, but in advocating translation strategies with the potential to foreclose identification with the foreign other, Venuti's theory valorises the Kantian impulse in the translator while distrusting it in readers. That is, Venuti's theory allows an ethics of alterity to guide the translator's choices, but does not appreciate the possibility that this is what guides some readers to read literature in translation in the first place. At the same time, we might also draw attention to the question of the appropriateness of using so-called "minority elements" (Venuti 1998: 11) to present a foreign culture or text. *If the problem discussed above is one of overzealously guarding access to the (foreign) cultural other, then here it is (partly, anyway) a problem of presuming access to the linguistic identities of domestic others in the service of that goal.* It is not difficult to see how such a

strategy might be construed as the translator appropriating minority cultural innovations.

While Venuti has taken pains to demonstrate how "foreignizing can cross the cultural boundaries between elite and popular readerships" ([1995] 2008: 121; see also Venuti 1998: 12), it no longer seems tenable to maintain a hard and fast distinction between these readerships, particularly when considered as distinct groups of people. As Boase-Beier (2020: 45) writes:

> […] the various types of reader proposed, either by Diaz-Diocaretz (1985: 15-33), or types of readers of translations, such as Venuti's "elite" and "popular" readers […], cannot really be considered to be types of people, but only ways of reading, tendencies, which might be exhibited more commonly by one reader than another, but that can in principle co-exist in the same reader.

By way of illustration, she appeals to the importance of the contexts in which readings take place: "I might be re-reading *Wuthering Heights* on the train, I might be re-reading it in order to teach a class on 'Mind, Body and Literature,' I might be about to begin a new translation into German, or I might be about to review a new translation" (ibid.). The crucial point here is: "Each of these [contexts] requires one to read differently" (ibid.). Thus, while Venuti (1998: 25) takes "scientific models" of translation research to task for ignoring the "cultural and social formations" in which linguistic practices take place, the same criticism could be levied at his conceptualisation of readerships. His recognition of "diverse cultural constituencies" (Venuti 1998: 9) as potential audiences for translations might seem like the opposite of a tendency to **universalise**, but if these constituencies are only thought to read in one of two modes, then diversity of response is severely artificially circumscribed.

***

In this section, I have identified and discussed some substantial similarities between Russian Formalism and Venuti's theory of foreignisation. These not only demonstrate the continued relevance of Formalist thinking to contemporary approaches to the reception of literature and literature in translation, but they also go some way toward establishing a grounds for empirically delineating *ostranenie* from foreignisation (if indeed they can be disentangled at the level of measurable effects). While there is ample precedent for using empirical methods to investigate FG, the amenability of foreignisation to empirical testing has yet to be demonstrated, and Venuti himself has argued against the "scientific model" of translation research (see e.g. Venuti 1998: 25-30; Venuti [1995] 2018: ix-x). Against Venuti's objections, I demonstrate in the next chapter that his claims are indeed amenable to empirical testing. In addition to discussing the specifics of my methodology in more detail, I also draw upon work by Umberto Eco to introduce a useful new concept—*intentio translatoris* [the intention of the translator]—to schematise the dynamics of reader response in Venuti's theory of foreignisation.

3. METHODOLOGY

3.1. OPERATIONALISING FOREIGNISATION

3.1.1. BACKGROUND: ECO'S *INTENTIO*S AND *INTENTIO TRANSLATORIS*

In order to test my hypotheses related to foreignisation, it was necessary to find some Japanese poems that had been translated at least twice into English, and to select among those 'competing' translations a pair for each ST such that one TT could be described as 'ostensibly foreignising' and the other 'ostensibly domesticating' based on textual criteria discussed by Venuti. (H_9 involved producing my own translation of a particular text, so I only needed one pre-existing TT to comprise a pair in that case.) The use of the word 'ostensibly' here is a response to the shifting criteria, discussed in the previous chapter, for the application of the terms 'foreignising' and 'domesticating' in Venuti's theory. In short: in some cases textual features seem to be the decisive factor in Venuti's theory, while in other cases it is the quality of attention paid by the reader that determines whether a text is foreignising; finally, in one instance, Venuti even refers to Burton's (1885/1888) orientalising translation of the Arabian Nights as "foreignizing in intention" (Venuti [1995] 2008: 269).

In this respect, Venuti's theory resembles an uneasy amalgamation of the three different approaches to interpretation to which Eco (1990: 47) argues "the whole history of aesthetics can be traced […]." According to Eco (ibid.: 50-51), three principles have historically been used to explain how texts produce responses in readers. These three principles are:

- *intentio auctoris* [intention of the author]

- *intentio operis* [intention of the text], and

- *intentio lectoris* [intention of the reader] (ibid.: 50).

As Miranda (1997: 372) summarises, "*intentio auctoris* assumes that interpreting a text means bringing to life the meaning the author had intended, and it assumes the objective nature of the text, independent of our own interpretations; whereas *intentio lectoris* assumes that there is an infinite number of interpretations of the text." Eco acknowledges that recourse to authorial intention[23] is not necessary to produce acceptable interpretations, yet he does not believe that all interpretations are equally acceptable, so there must be a constraining force at work. Whereas for Fish (1980b) acceptability depended upon the values of the interpretive community to which a reader belonged, for Eco (1990) constraints on acceptability are to be found in the text itself. Thus he appeals to the notion of *intentio operis*—in the broadest terms, this refers to "what the text says independently of the intentions of its author" (Eco 1990: 50). Eco's use of the terminology of 'intentio' and 'intention' may seem a bit puzzling at first glance. After all, as Pisanty (2017: 147) asks, "How can a text intend anything at all?" For Pisanty (ibid.), the use of 'intention' is a rhetorical device to personify the text as a "subject of volition." However, when we take that device away, what remains is the following hypothesis: "a text exerts a good deal of control over the range of its possible effects" (ibid.). In Eco's theory, this is a consequence of the fact that content words in a text have a semantic starting point that is difficult to ignore. To adapt one of Eco's examples, the word 'fig' has an everyday definition ('a type of fruit') that most speakers would be familiar with. This does not mean that it cannot take on additional meanings in a text, but he considers this first level of meaning to be a kind of constraint that "[n]o reader-oriented theory can avoid" (Eco 1990: 6). Nevertheless, if texts can produce effects independent of *intentio auctoris*, and these effects can be validated by appealing to the relative stability of the

---

[23] To clarify my own position on this issue, for the purposes of this study I have considered authorial intention as a kind of 'concept-by-postulation' (see below). That is, it is something that may be theorised to exist, yet which no one has ever directly observed: researchers have only ever observed things which may be argued to reflect intention (or not).

"material level of the text" (Easthope 1991: 38),[24] then it will be useful to consider why it is that in Venuti's theory, the three *intentio*s seem to be in such flux; and why *intentio operis* is itself not enough to validate readers' responses, as it is in Eco's theory. As I will argue below, the importance Venuti places on the stated aims of the translator—an issue I discuss under the heading of *intentio translatoris*—is a response not only to 'the translator's invisibility' but is itself the symptom of an instrumentalism that places its own constraints on the acceptability of responses.

If we consider a conventional reading scenario, wherein a reader reads a text written in a language they know well, it is clear that certain combinations of words are more likely to be encountered than others, as studies of colligation and collocation attest. This is a matter of linguistic probabilities of which we have unconscious knowledge. However, when we think back to Eco's (1990: 50) definition of *intentio operis*, it is notable that 'what the text says' is considered independently only from *intentio auctoris*—not *intentio lectoris*. As Eco (ibid.: 58) himself acknowledges, "it is possible to speak of text intention only as the result of a *conjecture* on the part of the reader." In other words, "[o]ne has to decide to 'see' it" (ibid.). In this sense, the construction of *intentio lectoris* resembles Venuti's 'symptomatic reading'; inasmuch as symptomatic reading "can be said to foreignize a domesticating translation by showing where it is discontinuous" (Venuti [1995] 2008: 24), both approaches emphasise "what the text says independently of the intentions" (Eco 1990: 50) of its author or translator, as the case may be.

In symptomatic reading, then, the reader is encouraged to become a kind of ethical (re)translator, and this repositioning is portrayed as attainable even to those readers who do not use the SL (see Venuti [1995] 2008: 21). As Venuti (ibid.: 21) writes, the "scientistic

---

[24] This refers to the idea that a text has a "relatively fixed identity" (Easthope 1991: 33), not least of all in terms of "the differential phonemic system of a language"; that is, any particular text will consist in "*these signifiers in this syntagmatic order and not others*" (ibid.).

strategy" of Freud's translators "can be demonstrated with no more than a careful reading of the English text." In paying attention to the diction of the TT, he identifies "an inconsistency in word choice which exposes the translation process" (ibid.: 22). Namely, he points out a sudden switch from "simple and common" expressions like "forgetting" and "go out of my head" to the more scientific "parapraxis" in a typical passage (ibid.). Implicit in this criticism is that the English word choice *ought to have been consistent*, owing to "Freud's heavy reliance on anecdotal, 'everyday' examples," as well as a footnote in a German edition emphasising that the work was self-consciously popular in tone. Thus, we see that when it comes to a "domesticating practice" such as the one employed in "the translations of the *Standard Edition*" of Freud (ibid.: 23), the *intentio translatoris* is *de-emphasised* in favour of what Venuti sees as the ST *intentio operis* and *intentio auctoris*. However, this is not so when it comes to foreignising translations (and particularly, his own translations), as we will see in a moment.

As Eco (1992: 43) writes, "If there is something to be interpreted, the interpretation must speak of something which must be found somewhere, and in some way respected." But how can the 'something' that precedes interpretation be strictly textual if it is posited in part by the reader? In fact, the two notions of *intentio operis* and *intentio lectoris* are problematically intertwined. As Eco (1990: 58-59) writes, "the intention of the text is basically to produce a Model Reader able to make conjectures about it"; meanwhile, the conjectures of this Model Reader are aimed at "figuring out a Model Author that is not the empirical one and that, at the end, coincides with the intention of the text." However, if the *intentio operis* is "the product of the process of interpretation," as Robey (2004: 8) remarks, then it "must result from the criteria that the process [of interpretation] projects into it." In other words, the *intentio operis* cannot be something essential or immanent to the text itself if, at the same time, it is a product of *intentio lectoris* (and therefore something that changes depending on the reader).

Nevertheless, Eco (1990: 58) insists upon maintaining this "dialectical link between *intentio operis* and *intentio lectoris*," while privileging the former as an explanatory principle for the effects of texts on readers. Moreover, when we consider that a reader's understanding of authorial intention can never be based on direct access to that intention—but rather on something else taken to be reflective of that intention, i.e. another text for the reader to interpret—it becomes clear that each *intentio* potentially informs the two others, as we can see in the diagram below. Insofar as Eco admits that *intentio auctoris* is not necessary to produce an interpretation, I have placed it in parentheses.

Fig. 3. Eco's *Intentio*s

(Intentio Auctoris)

Intentio Lectoris   Intentio Operis
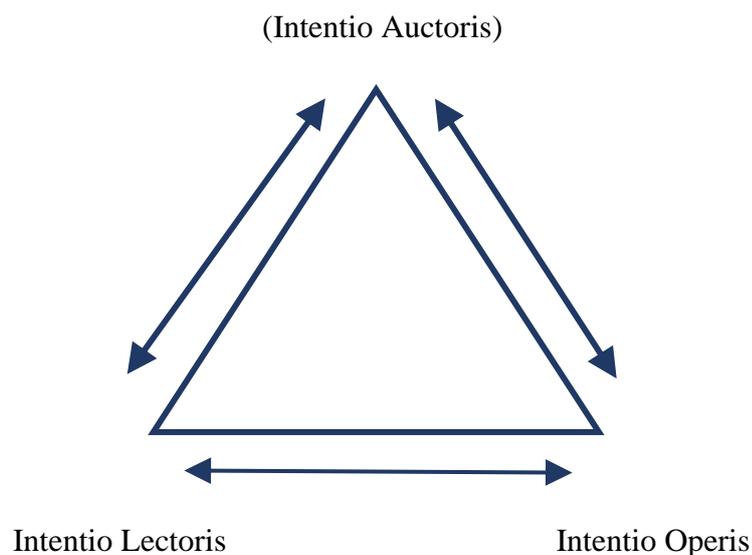
While there does not appear to be a simple solution to the paradox of an immanent *intentio operis*, Eco's theory is useful to think about in the context of operationalising foreignisation in my surveys. As I explain below, it helps justify my decision to categorise texts as 'ostensibly foreignising' or 'ostensibly domesticating' on the basis of textual features (with reference to reception scenarios).

Firstly, Venuti rejects the notion of the foreign as an "essence that resides in the foreign text" ([1995] 2008: 15). In this sense, the concept of the foreign does not seem to belong to the realm of the ST *intentio operis*. Rather, he describes the foreign as a "strategic construction whose value is contingent on the current situation in the receiving culture" (ibid.). In statements like this, Venuti effectively replaces *intentio auctoris* with *intentio translatoris*, and brings all three *intentio*s into play to account for the effects of foreignisation.

> The terms "domestication" and "foreignization" indicate fundamentally *ethical* attitudes toward a foreign text and culture, ethical effects produced by the choice of a text for translation and by the strategy devised to translate it, whereas terms like 'fluency' and 'resistancy' indicate fundamentally *discursive* features of translation strategies in relation to the reader's cognitive processing.

> (Venuti [1995] 2008: 19)

My argument here is that phrases like "the strategy devised" trigger a conventional implicature: they presuppose that *someone has intentionally come up with a plan about how to translate*. In every instance in Venuti's book, the person who has devised and implemented the translation strategy for a given TT is the translator (though they may have been influenced by their institutions, clients, other writers, theorists, etc.). This would seem to augur for *intentio translatoris* as the explanatory principle behind the effects of foreignisation.

This seems the case particularly when he discusses the reception of his own translation work: citing a rather negative review of his translation of I.U. Tarchetti's novel *Passion*, Venuti (1998: 19) writes that the reviewer "refused to understand" the effect he was aiming for "according to the explanation presented in my introduction […]." There is a sense in

which, when the translator has a foreignising intention in Venuti's model, the *intentio translatoris* is supposed to be respected. Thus, when a translation does not have the desired effect, one may point to the idea that the reader did not read it appropriately—it is a faulty *intentio lectoris* that is to blame. In this sense, Venuti's theory is open to the same criticism he levies against the instrumental model of translation, in which translation is conceived as the "reproduction or transfer of an invariant that is contained in or caused by the source text, an invariant form, meaning, or effect" (Venuti 2019: 1). The difference here is: it is not the ST that is treated as if it contained an invariant, but the TT. The foreignising *intentio translatoris* is itself treated as a kind of invariant—something supposed to guarantee an effect, such that when the said effect is *not* produced, it can only be the case that a particular reader has not properly understood the *intentio translatoris*.

Venuti's model is similar to Eco's in the sense that *intentio translatoris*—like *intentio auctoris*—blinks on and off, as it were, as a component in the diagram below. However, whereas for Eco, recourse to *intentio auctoris* is a kind of optional step in the reader response process, in Venuti's theory recourse to *intentio translatoris* sometimes *is* absolutely necessary in order for readers to respond in the appropriate way—to prevent the collapse of intention and effect into each other, as we saw in Russian Formalism. That is, recourse to *intentio translatoris* is sometimes necessary to avoid the conundrum of a translation which is foreignising in intention but not in effect: Venuti's theory assumes that knowledge of such intention (correctly apprehended) would either guarantee that readers responded in the appropriate way (i.e., experienced the effects of foreignisation), or at least show them how their initial responses were wrong.

Fig. 4. ST and TT *Intentio*s

ST   TT

(Intentio Auctoris)   (Intentio Translatoris)

A

Intentio Operis$_1$ Intentio Lectoris$_1$ Intentio Operis$_2$ Intentio Lectoris$_2$
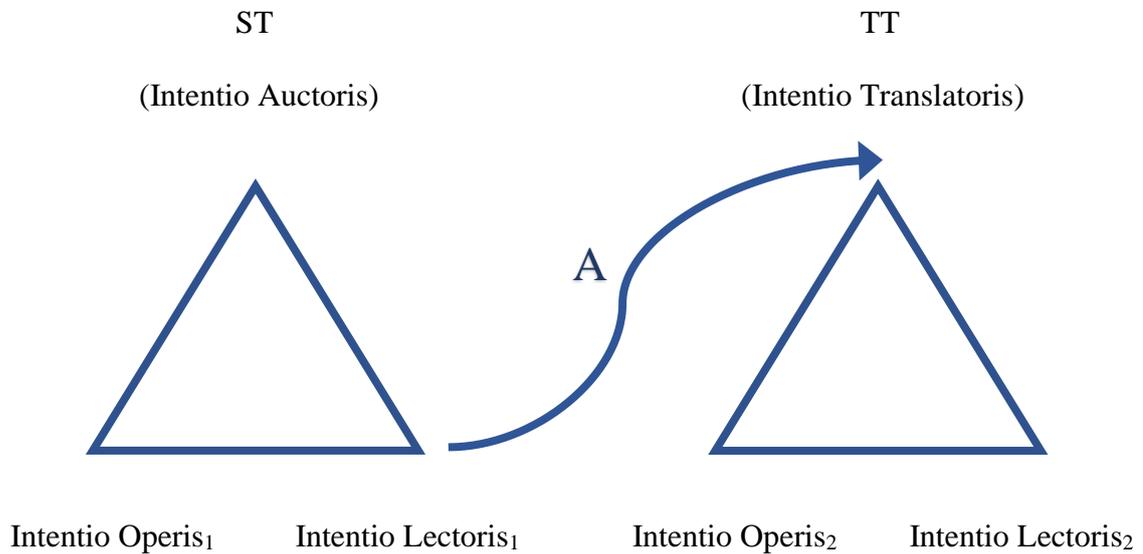
Moreover, Venuti's emphasis on the illusory aspect of authorial presence in the translation would seem to strengthen the disjunction between ST and TT, such that the translator arguably takes the author's place in a similar triangle of *intentio*s on the TT side of the diagram. However, the disjunction is not total; the ST triangle and the TT triangle are linked by a kind of intermediate *intentio*—one that is "neither reader nor writer but somehow both"[25]—bridging the ST *intentio lectoris* position and the TT *intentio translatoris* position (this is the line labelled 'A').

And yet, Venuti also acknowledges that not all translation choices can be described as conscious or intentional parts of a strategy, as we see in his discussion of the "abusive fidelity" of I.U. Tarchetti's translation/plagiarism of Mary Shelley's "The Mortal Immortal" (ibid.: 150). Shelley's tale contains a reference to a popular Orientalist text by Frances Sheridan entitled *The History of Nourjahad* (1767), a reference which Tarchetti deleted in his

---

[25] Thanks to Dr Caroline Summers for suggesting this phrase.

1865 translation (published under his own name, with no mention of Shelley's authorship). Venuti argues that while Tarchetti's deletion (not to mention his plagiarism) "points to an antifeminist effect in the translation" (ibid.: 149), it also performs an "ideological critique of Shelley's tale, exposing the political limitations of her feminism, its failure to recognize the gender hierarchy in the bourgeois marriage and its concealment of working-class oppression and European racism" (ibid.: 150). Venuti cannot say for certain what Tarchetti's motive was in removing Shelley's allusion—he supposes that it "may have been merely due to his ignorance of Sheridan's tale" (ibid.: 149). But to continue to hold Tarchetti's text up as an example of foreignising translation, he needs to point to something beyond *intentio translatoris* in this instance. Thus, he relies on the notion of *intentio operis*, i.e., "what the text says independently of the intentions of its author" (Eco 1990: 50)—or what the text says independently of the intentions of its translator, in this case. As Venuti (1995 [2008: 150]) writes, "the foregoing treatment of Tarchetti's translation requires a revision of [abusive fidelity] to include translation choices that remain unarticulated and unconscious, and that therefore can support an effect exceeding the translator's intention."

Since the "'foreign' in foreignizing translation is not a transparent representation of an essence that resides in the foreign text" but a "strategic construction" on the part of the translator (Venuti [1995] 2008: 15), it is clear that in Venuti's theory, the author has no power to foreignise (unless perhaps the author is also the translator of their own work). Moreover, as he suggests in his discussion of Schleiermacher's work, any sense of the "presence of the foreign author in the translation" is only "illusory" (ibid.: 95). Thus, we are dealing with a model of response in which the effects of foreignisation are decidedly associated with the TT side of the ST/TT dichotomy. But what does this imply about the presence of the translator?

Discussing John Nott's (1795) foreignising translation of Catullus, Venuti writes that

Nott's "main concern seems to have been twofold: to ward against an ethnocentric response to the Latin text and to signal its historical and cultural differences" (ibid.: 71). Venuti goes on to discuss details of the translator's life that make this imputation of intention at least somewhat plausible. But again we seem to be in the realm of phenomena that readers must decide to see, to paraphrase Eco again. This seems to be down to *intentio lectoris* again. However, whereas in symptomatic reading, the reader was obliged to become a kind of (re)translator, here they must become a historian of translation. But what if a translator set out with the opposite intention—with the intention to domesticate in wholesale fashion? And what if their translation nevertheless registered for readers as foreignising? Should we describe that translation as foreignising, despite the evident lack of a foreignising strategy or intention? Venuti's repeated use of contemporary reviews to gauge the effects of foreignisation, along with his amendment of Lewis's (1985) 'abusive fidelity' to encompass effects that exceed intention, seem to suggest an answer: yes.

In Venuti's theory, as in Eco's, a text cannot have an effect without a reader. Sometimes Venuti's theory demands much of readers in order to remain coherent—I refer here to readers who do not automatically respond in the desired fashion, or who are otherwise forced to foreignise a domesticating translation themselves via symptomatic reading. Other times, the theory requires less of them. Indeed, in most cases discussed in *The Translator's Invisibility*, the effects of texts are gauged with reference to the reactions of contemporary reviewers—or, failing that, they are posited on the basis of educated conjectures about hypothetical readers. For instance, the "potential meanings" of Shelley's reference to *The History of Nourjahad* "*would have been* accessible to readers of *The Keepsake*," the literary annual that first published "The Mortal Immortal" (Venuti [1995] 2008: 149; emphasis added). But behind both kinds of evidence, the dynamic of norm-breaking looms large. As Venuti (ibid.: 85) writes, quoting Schleiermacher, "the 'innovations and deviations' of foreignizing translation

are defined against the norm set by other translation discourses in the receiving culture." But it is not just other translations that a foreignising translation might deviate from; it is the canon of non-translated literature as well. Venuti (ibid.: 153) again: "In the case of prose fiction, the translator can select a foreign text whose fictional discourse or genre runs counter to the narrative forms that have achieved canonical status in the literary traditions of the translating language." This would seem to hold for poetry, too. But just what does a challenge to a canon look like? Determining the canonical literary values of the present day is tricky: first, which canon are we talking about? The canon of translated Japanese literature in translation? Or canonical forms of non-translated English-language poetry? In my analyses of the stimulus texts that follow, I try to strike a balance between the broader backdrop suggested by the Royal Society of Literature's 2017 readership survey "Literature in Britain Today" and the narrower backdrop of the reception of Japanese literature in the UK.

Nevertheless, the preceding discussion has important ramifications for the survey design. As the emphasis on norm-breaking suggests, a translation may be described as foreignising in Venuti's theory without recourse to *intentio translatoris*. To the extent that the breaking of norms involves specific textual features—and in order for a norm to be broken, it must be recognised by a reader as a norm (however unconsciously)—the two explanatory principles here would be *intentio operis* and *intentio lectoris*. As none of the respondents in my surveys were provided with information about the translators, and it is unlikely that they sought out such information before they filled out the surveys, this helps justify my decision to categorise the stimulus texts on the basis of textual features (with reference to reception scenarios). In other words, we cannot attribute foreignising effects directly to readers' research into translator intention. (Of course, this would not prevent readers from forming their own opinions about the intentions of the translators.) It could be argued that familiarity with the theory of foreignisation might allow some readers to perform a symptomatic reading

of the texts, producing their own foreignising effects. However, I included a survey question about familiarity with the theory in order to control for that variable. (Of course, just because readers are unfamiliar with the concept as it is named does not mean that they are unfamiliar with the concept as such. However, this will have to remain a limitation of the current study.)

   In addition to justifying the decision to categorise texts in this fashion, this discussion has uncovered several consequences of Venuti's theory not addressed in previous research. In cases like Burton's translation of the *Arabian Nights* where Venuti appeals directly to *intentio translatoris*, as well as in other cases where Venuti ascribes to the translator a will-to-foreignise, the TT becomes a kind of conduit for an invariant effect, so that Venuti's model does not really escape instrumentalism after all. If instrumentalism "conceives of translation as the reproduction

or transfer of an invariant that is contained in or caused by the source text, an invariant form, meaning, or effect" (Venuti 2019: 1), then here the invariant is caused by the foreignising translator, such that when the effects of foreignisation are *not* felt by readers, this is seen as simply an inappropriate response on their part. But that is not all. Venuti does not really discuss how the effects of foreignisation vary *within those populations in which they are attested*, nor does he delve too deeply into *how* those effects on readers contribute to more equitable working conditions for translators or the "emergence of new cultural relations" (Venuti 1995 [2008]: 277). This is especially significant given the possibility that those readers most responsive to foreignising translations might already be predisposed to recognise and value cultural differences. There is a sense in which the knock-on effects of foreignisation are taken for granted. This is an extension of the invariance described above. Venuti (2019: 1) might be opposed to viewing translation as the "reproduction or transfer of an invariant that is contained in or caused by the source text," but in his reliance on *intentio translatoris* (by which he reserves the right to characterise a translation as foreignising on the

basis of research into the translator's aims), his theory is capable of treating foreignising translation as the transfer of an invariant effect that is caused by the translator. The fact that he does not recognise this, even in his most recent work, suggests a place for *intentio translatoris* in future theorising about how readers respond to translated literature.

More than this, the notion of *intentio translatoris* reveals the difficulty of escaping the hermeneutics of freedom and fidelity traditionally associated with the study of translation. *While Venuti's model problematises the notion of any direct, horizontal transfer between ST and TT, it nonetheless seeks to establish a direct, vertical transfer of effect from translator to reader in the TT reception scenario*. Moreover, the notion of a correct response recalls Eco's notion of the Model Reader, capable of discerning the *intentio operis* and responding accordingly; but it also hints at the underlying sympathy or accord that must apparently exist between translator and reader in Venuti's theory for the reader to respond in the *right* manner. This is a consequence of the fidelity/freedom hermeneutic. As Frölich (2017: 36) writes, "Behind this conception, that the interpreter repeats the process of forming the artwork in his own way, lies the assumption that a congeniality exists between artist and interpreter […]." Only in this case, we substitute 'reader' for 'interpreter,' and 'translator' for 'artist.'

Finally, apart from symptomatic reading, where the reader seems to reign supreme, the translator is ironically granted more freedom of response than other readers in Venuti's theory. In fact, it is thanks to the freedom inherent in *intentio lectoris* at the point of encounter with the ST that *the translator is able to assume a position similar to the auctoris* when it comes to the TT in Venuti's theory. What I mean is: Venuti's disjunction between the ST and TT underwrites the freedom of the translator *as reader of the ST* to insist on their own interpretations and even contravene what they understand to be the ST's *intentio auctoris* in order to translate in a liberatory fashion. In other words, if the translator becomes (something like) the *auctoris* of the TT, they are free to insist on their own interpretation of the ST, and

as long as their *intentio translatoris* is felt to issue a challenge to (certain) dominant values in the translating language and culture, the burden is on the reader to align themselves with this *intentio* to be able to respond in an 'acceptable' fashion. In this sense, the translator-as-reader-of-ST is the party with the most freedom: despite Venuti's characterisation of foreignising translation as "beset with risks" owing to rigid "[c]anons of accuracy" in "contemporary British and American cultures" (Venuti 1995 [2008]: 273), it is the translator who ultimately sets the interpretive and ethical agenda in the scheme outlined above. Readers who respond negatively are treated as a hindrance to cultural change, while readers who respond positively do so in an almost passive way that is assumed to contribute to greater intercultural understanding and the material improvement of translators' conditions. How much of the ST's *intentio auctoris* survives in translation? Or to pose the question another way: to what extent is the *intentio translatoris* a function of 'canons of accuracy' in the translating scenario? These are questions which need to be addressed to move the debate forward. But before we tackle those questions, it would be useful to test the assumption that foreignisation actually produces the effects claimed for it.

3.1.2. TEXTUAL AND RECEPTION CRITERIA

The textual criteria for foreignisation that I focus on derive from two separate but related domains in Venuti's theory. The first is his general theorising on foreignisation. This is broadly concerned with the instrumental or conduit model of translation; to use Venuti's terminology, this is the notion of translation as a "transparent discourse" akin to a window onto the foreign text, an idea that has "dominated English-language translation since the seventeenth century" (Venuti [1995] 2008: 177). And the second domain is his discussions of foreignisation in the specific context of Japanese-to-English literary translation. While I have

called the former level of theory 'general,' it is clear the Venuti has arrived at it through analysis of specific TTs from specific language pairs (though largely translations *into* English). Where *The Translator's Invisibility* is concerned, Venuti has decried the "tendency to emphasize the overall account of the project in the first chapter" at the expense of the "case studies in the remaining chapters" (Venuti [1995] 2018: xi). Venuti objects to a (perceived) lack of attention paid to the ways in which cultural and historical contexts affect the phenomena he has outlined in the first chapter; he emphasises that it is the case studies later in the book which demonstrate how such concepts as foreignisation and domestication "change in specific cultural situations at specific historical moments" (Venuti [1995] 2018: xi). Yet despite his avowed scepticism of "scientific models" of translation research (see, e.g., Venuti 1998: 25-30), Venuti's work on foreignisation can indeed be described in terms of the traditional scientific research paradigm. This paradigm was famously diagrammed by Walter Wallace in *The Logic of Science in Sociology* (1971: 18), where it takes the form of a wheel whose components inform one another in a continuous process; moreover, this paradigm involves two kinds of logic (deductive and inductive), as we can see in the adapted diagram below:

Fig. 5. The 'Wheel of Science'
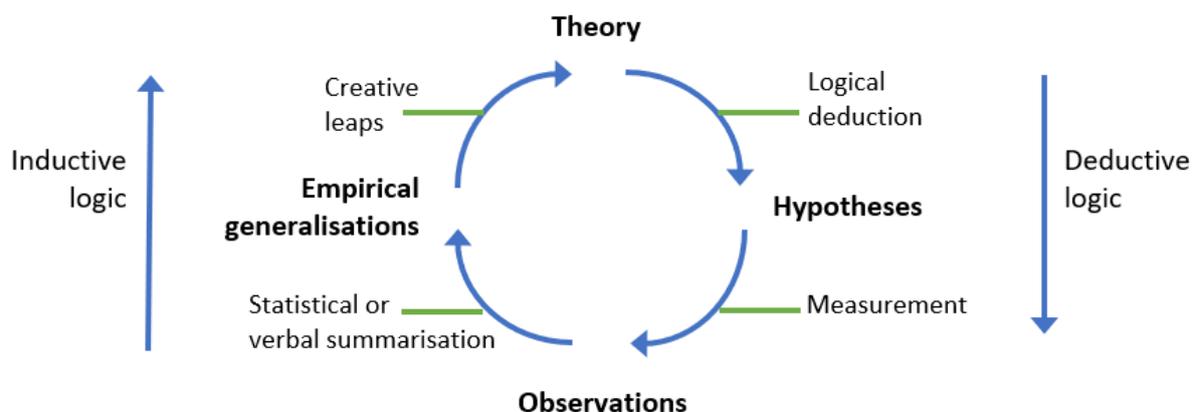
(adapted from Adler and Clark 2015: 33; originally adapted from Wallace 1971: 18)

Venuti may not organise or present his work like a scientific study, but this does not mean that his research does not generally follow the logic outlined above. What Venuti presents in the introductory chapter of *The Translator's Invisibility* is no less than the outcome of the inductive side of the above 'wheel of science,' from 6 o'clock (Observations) to 12 o'clock (Theory). In other words, starting from specific observations about the reception of fiction translated into English,[26] he generalises that "fluency" is the dominant criterion by which such translations are currently judged (Venuti [1995] 2008: 2). Moreover, he reaches this stage of generalisation via verbal summarisation, just as in the 'wheel of science' diagram. That is, he summarises his observations: "The critical lexicon of literary journalism since World War II is filled with so many terms to indicate the presence or absence of a fluent translation strategy […]" (ibid.: 4). And he generalises from that summary: "A translated text, whether prose or poetry, fiction or nonfiction, is judged acceptable by most publishers, reviewers and readers when it reads fluently […]" (ibid.: 1). We note here that Venuti seems to have used observations concerning one text type (fiction) to make a generalisation about multiple text types; indeed, this is one place where the organisation of his book obscures the logical process at work: it soon becomes clear that he has 'observed' not just the reception of modern translated fiction, but the reception of many different kinds of texts in translation: there are tales, poetry, drama and the Bible, not to mention works of biography and psychology. It is clear that his observations about these text-types have *also* informed the generalisation cited above; they are just presented later. And this generalisation—that fluency is the main criterion for the acceptability of translations—is augmented by another. From his observations, Venuti also generalises about what constitutes fluent translation: this kind of translation is produced by "**adhering to current usage,**

---

[26] The observations come from reviews of fiction translated into English from a variety (read: sample) of British and American publications.

maintaining continuous syntax, [and] fixing a precise meaning" (Venuti [1995] 2008: 1; emphasis added). More specifically, such translation is "**written in English that is current ('modern') instead of archaic, that is widely used instead of specialized ('jargonization'), and that is standard instead of colloquial ('slangy')**" (ibid.: 4; emphasis added). Moreover, in a fluent translation "**foreign words or English words and phrases imprinted by a foreign language ('pidgin') are avoided, as are Britishisms in American translations and Americanisms in British translations**" (ibid.; emphasis added). Additionally, fluent translation "**also depends on syntax that is not so 'faithful' to the foreign text as to be 'not quite idiomatic,' that unfolds continuously and easily ('breezes right along' instead of being 'doughy') to insure semantic 'precision' […]**" (ibid.; emphasis added). Finally, fluent translation is "**immediately recognizable and intelligible, 'familiarised,' domesticated, not 'disconcerting[ly]' foreign, capable of giving the reader unobstructed 'access to great thoughts,' to what is 'present in the original'**" (ibid.: 4-5; emphasis added).

I understand from Venuti's complaint in the third edition of his book (Venuti [1995] 2018: xi) that what is considered a fluent translation may change depending on the cultural and historical situation, and that the above description of fluency holds only at the very general level of "contemporary Anglo-American culture" (ibid.: viii). While the features above helped me categorise the competing translations in my survey, before we discuss the translations in detail, we need to understand the link between fluency and domestication/foreignisation. As it turns out, Venuti establishes this link via another step in the 'wheel of science' diagram. I am referring here to what Adler and Clark (2015: 33) call "creative leaps," the intermediate step that transforms empirical generalization into theory. According to Wallace (1971: 17), one way that empirical generalisations are "synthesized into a theory" is through "concept formation." There are two possible procedures for concept

formation: the first is by "making the terms and relationships in empirical generalisations more abstract," and the second is by "introducing other abstract terms that refer to nonobservable constructs" (ibid.: 53). What results is a theory: a "structure" that "can *explain* known empirical generalizations" and which "can *predict* empirical generalizations that are still unknown" (ibid.: 57).

In Venuti's case, the notion that fluent translations are more acceptable than non-fluent translations is made more abstract through a similar process. It is not just that fluent translations are more acceptable; it is that they are more acceptable *because* they produce an effect known as "transparency" (Venuti [1995] 2008: 1), and they produce this effect in a contemporary reception scenario that generally values it (however broadly defined that scenario might be in Venuti's work). By 'transparency,' Venuti (ibid.: 47; emphasis added) refers to an "illusionistic effect" by which "**the translation seems as if it were not in fact a translation, but a text originally written in English**." Moreover, there are several reasons why transparency "is the dominant discourse in poetry and prose, fiction and non-fiction, best-sellers and print journalism" (ibid.: 97). One of these reasons, according to Venuti, is the long-standing influence of instrumental views of language (such as the conduit metaphor described by Easthope 1983: 11; see also Venuti [1995] 2008: 49). Another reason is economic in nature: the consumability of transparent texts in the "contemporary cultural marketplace" (ibid.: 97) seems to drive the production of *more* transparent texts. Thus, in the process of concept formation, Venuti moves from the observable realm of correlation to the more abstract realm of causation (see e.g. Wallace 1971: 27). In other words, the correlation between fluency and acceptability is something made on the basis of observables: the number of reviews praising specific linguistic features in relation to fluency. Meanwhile, the relationship between fluency and acceptability is *causally* explained by the addition of a non-observable concept ('transparency').

The next task in constructing a theory (we are still in the 'creative leaps' area of the diagram) is what Wallace (1971: 53) calls "proposition formation." Wallace (ibid.: 55) writes: "Once the required and appropriate observables and unobservables have been conceptualized, propositions are formed by fitting them into the form, 'If concept X, then concept Y,' or 'The greater the X, the greater the Y.'" What we find on page 1 of *The Translator's Invisibility* is exactly this sort of statement: "The more fluent the translation, the more invisible the translator, and, presumably, the more visible the writer or meaning of the foreign text" (Venuti [1995] 2008: 1). In other words, the more fluency, the more transparency. Unfortunately for the researcher hoping to operationalise Venuti's theory, he has shied away from tendency statements like these when it comes to two other unobservable constructs: foreignisation and domestication. But this does not mean that his work does not repeatedly present certain configurations of fluency, resistancy, domestication and foreignisation as typical of Anglo-American reception contexts. If Venuti has avoided calling attention to these typical configurations himself, it may be because he wanted to try to avoided the kind of "reductive misunderstandings" (Venuti 1995 [2008]: ix) of his concepts which he decries in the second and third editions of his book. One of these misunderstandings has been to associate domestication and foreignization with fluent and resistant "discursive strategies" respectively (ibid.: 19). And yet, this emerges as an understandable assumption when one does look closely, as Venuti urges, at the case studies in his work. In fact, apart from one (arguable) instance,[27] every translation or translation practice described as somehow resistant is associated with foreignisation. As Venuti (ibid.: 266) puts it: "Resistance assumes an ethics of foreignization […]." Moreover, every translation or translation practice described as domesticating is described as such in the immediate or extended context of transparency.

---

[27] This is William Arrowsmith's translations of Montale. Venuti ([1995] 2008: 247) writes that the "modernist translation discourse Arrowsmith recommended may have been resistant to certain British and American values […], but his rationale for this discourse agreed with mainstream poetics, the romantic valorization of the poet's 'voice.'" Thus, Venuti stops short of calling Arrowsmith's version 'resistant' *tout court*.

Where the latter is concerned, one would be forgiven for thinking that transparency is a necessary—if not necessary and sufficient—condition for applying the term 'domesticating' to a translation (in the subordinate sense of the term discussed previously). Likewise, outside of the warning that "foreignization" cannot "simply be superimposed on […] 'resistant' discursive strategies" (ibid.: 19), Venuti provides no cases of resistant translations that fail to be foreignising. Since we have to start somewhere, we would be remiss not to consider textual features associated with resistancy as potential predictors of foreignisation.

There is one element of Venuti's warning, however, that we would do well to heed. We know that in Venuti's theory, transparency is produced by fluency; moreover, Venuti describes all transparent translations as domesticating. However, not all fluent translations produce the illusion of transparency. This is especially clear when he discusses Japanese fiction in English translation. As I mentioned above, this is the second domain from which the textual criteria I discuss are derived. (Venuti discusses no Japanese poetry, only fiction. But again: we must start somewhere.) Venuti (1998: 85; emphasis added) describes Megan Backus's 1993 translation of Yoshimoto Banana's novel *Kitchen* thus:

> This version is highly readable, but it is also foreignizing in its translation strategy. Instead of cultivating a seamless fluency that invisibly inscribes American values in the text, Backus developed an extremely heterogeneous language that communicates the Americanization of Japan, but simultaneously foregrounds the differences between American and Japanese culture for an English-language reader. The translation generally adheres to the standard dialect of current English usage, but this is mixed with other dialects and discourses. **There is a rich strain of colloquialism, mostly American, both in the lexicon and the syntax** […]. **There is also a**

recurrent, slightly archaic formality […]. **There are, moreover, many italicized Japanese words scattered throughout the text** […].

Venuti revisits this translation in *The Translator's Invisibility*, where he writes that the reception of this translation, together with the reception of his own translations of Tarchetti, makes it "quite clear not only that foreignizing can cross the cultural boundaries between elite and popular readerships, but that it can change reading patterns by broadening the spectrum of linguistic forms used in translating and thereby redefining commonly accepted notion of fluency" (1995 [2008: 121]). While I have problematised the notion of elite and popular readerships above, there is one more thing to say before moving on to the analyses of the stimulus texts. In the wake of the extraordinary popularity of Murakami Haruki (whose works are known in part for frequent allusions to American popular culture), there is ample room to doubt that texts presenting "an Americanized Japan" (ibid.) would still run counter to popular perceptions of Japan.

3.1.3. STIMULUS TEXTS (OSTENSIBLY FOREIGNISING VS. OSTENSIBLY DOMESTICATING)

As I mentioned earlier, the use of the term 'ostensibly' is motivated by the shifting criteria by which Venuti has designated translations as foreignising. However, as I also demonstrated above, he has frequently associated foreignisation in Anglo-American reception scenarios with a consistent set of textual features (I bolded these in the previous section). Moreover, his theory allows for these features to have a foreignising effect on readers independently of *intentio translatoris*.

Nevertheless, while Venuti has argued that "**[a]ll translation** […] **is an interpretation that fundamentally domesticates the source text**" (Venuti [1995] 2018: xii), there is a

sense in which all the stimulus texts in this project may be said to "deviate from dominant literary canons in the receiving culture," to put it in Venuti's ([1995] 2008: 152) terms. It might seem outmoded to speak of a 'receiving culture' in the singular—we would probably want, ideally, to further specify *which* receiving culture(s) it is whose canons were being challenged. Nevertheless, interesting work has been done on the broad category of 'literature in Britain today.' In fact, this was the title of a 2017 survey conducted by the Royal Society of Literature. This survey was "carried out by Ipsos MORI between 23 September and 11 October 2016. Interviews were conducted face to face in home among a nationally representative sample of 1,998 adults aged 15+ in Great Britain (excluding Northern Ireland)" (Royal Society of Literature 2017a: 9). Respondents in this survey were asked, among other things, to name "a writer, living or from the past, whose work they would describe as literature" (ibid.: 23). The researchers report that "the 50 authors most frequently named" as writers of literature were "all white"; moreover, the only non-white writers cited by more than one respondent were Murakami Haruki and Zadie Smith (ibid.: 7). However, since the top 50 writers (Royal Society of Literature 2017b: 113ff) tend to be known primarily for their novelistic output (with a few exceptions like Shakespeare at #1, or Wordsworth at #22), it is probably more appropriate to compare the poets in my survey to those poets most frequently cited in the RSL survey. Excluding Shakespeare, the Brontë sisters and Thomas Hardy, who are assigned to other categories in the RSL survey, the researchers report that the following poets were named by more than one respondent:

1. William Wordsworth

2. T S Eliot

3. Robert Burns

4. Lord Byron

5. Geoffrey Chaucer, John Keats, Alexander Pushkin

8. Mihai Eminescu, Homer, Dylan Thomas

Once again, the list skews toward writers who wrote in some variety of English, although the proportion of writers who wrote in a language other than English is higher here than in the overall top 50. That said, there are certainly no poets from East Asia. In this sense alone, all of my stimulus texts could be described as presenting a challenge to the above canon. There are also no women in the above list: the three female poets in my survey could thus be described as presenting even more of a challenge. *In this sense, it might be more appropriate to discuss the 'competing' stimulus texts as 'ostensibly more foreignising' and 'ostensibly less foreignising' rather than use a category like 'ostensibly domesticating.'* But then, Venuti does not have much to say about the parameters by which we may judge or ascertain a challenge to a literary canon. Moreover, some of the stimulus texts *do* seem to have features that Venuti associates with fluency and domestication. In the end, I chose to stick with my original categories ('ostensibly foreignising' and 'ostensibly domesticating') because they are easier to tell apart at a glance. The tentativeness of the categories should be obvious from the use of 'ostensibly.' I discuss the competing translations—and their categorisation—below.

The first translation I discuss is Janine Beichman's (2007: 416) translation of Ishigaki Rin's "Shijimi" ([1968] 1998: 67). I have categorised this translation as **ostensibly foreignising**. To begin with, I encountered three other translations of this poem, each from a different decade: Kenneth Rexroth and Ikuko Atsumi's (1977: 96), Hiroaki Sato's (1981a: 574) and Leith Morton's (2005: n.p.). I present them for the sake of reference below:

Shijimi Clams

woke up in the dead of night--
in a corner of the kitchen
the little clams I'd bought that evening
were alive, mouths open--

"At dawn
I'll gobble you up
each and every one"

let out a cackle
like an evil old witch
after that couldn't help it had to
sleep all night with mouth half-open

(Beichman 2007: 416)

CLAMS

In the night I awoke.
The clams I bought yesterday
In a corner of the kitchen
With mouths open were alive.

'When dawn comes
I'm going to gobble them all up
Every single one.'

I cackled
The cackle of a witch.

From that moment on

My mouth slightly open

I passed the night in sleep.

(Morton 2005: n.p.)

Clams

At midnight I woke up.

The clams I'd bought in the evening

were alive in a corner of the kitchen,

their mouths open.

"In the morning

I'll eat you,

every last one of you."

I laughed a witch's laugh.

After that

I could only sleep through the night,

my mouth slightly open.

(Sato 1981a: 574)

Shellfish

I wake at midnight.

The little shellfish I bought last evening

are alive with their mouths slightly open.

I will eat them all when day breaks.

I laugh a hag's laugh.

Afterwards there is nothing left of the night,

except to sleep with my mouth slightly open.

(Rexroth and Atsumi 1977: 96)

However, only Beichman's retained any Japanese vocabulary from the ST (the word

"Shijimi.") Venuti might nowadays roll his eyes at this kind of understanding of his

theory; indeed he takes Matthew Reynolds (2016) to task for "reduc[ing] foreignizing

translation to a specific verbal choice or discursive strategy," in this case "literalism,"

i.e. "close or exact adherence to the source text" (Venuti [1995] 2018: xviii). However,

Venuti has indeed identified the retention of ST vocabulary as a factor in two

supposedly foreignising translations from the Japanese: first in Megan Backus' 1993

translation of the Yoshimoto Banana novel *Kitchen* (Venuti 1998: 85-86; Venuti [1995]

2008: 121), and then in Alfred Birnbaum's 1996 translation of Miyabe Miyuki's novel

All She Was Worth. The next thing that struck me as the kind of feature Venuti might

emphasise was Beichman's sparing use of "I," which recalls the so-called 'zero

pronoun' of Japanese syntax (Tsujimura 2014: 252). 'Zero pronoun' refers to a kind of

"[n]oun ellipsis whose interpretation is subject to contextual information shared among

speech participants" (ibid.). As Tsujimura (ibid.: 251) writes, English permits a "limited

degree" of this kind of ellipsis: for instance, a sentence like 'Bought a book yesterday' is

typically judged to be grammatical only in certain contexts, such as in casual emails or

text messages. Japanese, on the other hand, "allows noun ellipsis to a greater degree than

in English" (ibid.). Beichman's translation uses "I" only twice, compared to five

149

instances in Sato's translation, four in Rexroth and Atsumi's, and five in Morton's.

Where poetry translation is concerned, pre-empting or frustrating "sympathetic identification" ([1995] 2008: 211) with the lyrical 'I' is associated with strategies of resistancy and foreignisation in Venuti's work. For instance, remarking on a poem by Milo De Angelis, Venuti (ibid.: 250) writes that the "text does not offer a coherent position from which to understand it or a psychologically consistent voice with which to identify. On the contrary, the fragmented syntax and abrupt line-breaks constantly disrupt the signifying process, forcing the reader to revise his interpretations." Venuti writes that to translate the poem, he could have chosen a fluent strategy and thereby attempted to "mitigate the grammatical uneasiness usually provoked by the omission of a subject or verb in an English sentence"; however, his "English version […] refuses fluency" and "seeks to reproduce the discontinuity of De Angelis's poem" (ibid.: 251). Moreover, he decides that "the translation is no doubt more discontinuous with the omission of a subject […]" (ibid.). According to this explanation, we might also expect Beichman's text to frustrate readers' identification with the lyric voice—however briefly or intermittently—given that it uses the 'I' so sparingly. Not only that, but the use of dashes in her translation would seem recuperable to the category of 'abrupt line-breaks.' There is also the fact, discussed later in my stylistic analysis of the poem, that line 10 runs two sentences together, and it also does not terminate at a major constituent boundary, as the other lines tend to do. This seems the opposite of "maintaining continuous syntax," a feature Venuti ([1995] 2008: 1) associates with fluency and domestication. Furthermore, the type of ellipsis we see in Beichman's translation is also called "Main Subject Deletion" (Gandón-Chapela 2013: 295). As Gandón-Chapela demonstrates, this kind of ellipsis is most common in personal journals and diaries (ibid.: 304). Thus, at the same time that Beichman's translation may temporarily pre-

empt identification with the lyric voice, in certain respects the speech act that her text most resembles (outside of a poem) is a diary entry—one of the most personal forms of writing in English. Given that the speaker says they "let out a cackle / like an evil old witch," it does not seem too farfetched to expect that most readers would infer that the speaker is female. So: a woman's diary entry. To the extent that canons of English-language poetry have historically privileged men's voices and experiences—a picture largely supported by the RSL survey results—Beichman's text might reasonably be described as presenting a challenge to the gender values inscribed in such a canon (even if respondents to my survey do not know the author's gender.)

All the translations of this poem, apart from Morton's, use italics for what the ST (lines 6-7) has in the *katakana* syllabary,[28] where the speaker indicates a plan to eat the clams. But the final factor that led me to choose Beichman's translation as the ostensibly foreignising one was the shift in discourse used to translate that passage: 'At dawn / *I'll gobble you up / each and every one*.' Heterogeneity of discourse is associated with foreignisation in Venuti's theory ([1995] 2008: 28-29). The use of the word 'gobble' (also seen in Morton's translation) instead of 'eat' recalls the language of fairy tales; meanwhile the use of the adjunct 'each and every one' at the end of the sentence to qualify the preceding object 'you' might even put some readers in mind of the famous threat of the Wicked Witch of the West: "I'll get you, my pretty, and your little dog, too!" Both Sato's and Morton's translations employ an adjunct construction similar to

---

[28] As Perea, Nakayama and Lupker (2017: 1140-1141) explain: "In addition to its logographic Kanji characters, Japanese has two Kana syllabaries, Katakana and Hiragana, with both of these syllabaries having characters with direct correspondences to Japanese syllables […]." However, the authors continue, "the use of the two Japanese scripts follows precise rules. On the one hand, Hiragana is usually used to complement Kanji words (e.g., postpositional particles […] [and] inflections […]) although it can also be used transcribe some content words. On the other hand, Katakana is used primarily to transcribe loan words […]" (ibid.: 1141). It is also frequently used for sound effects and exclamations in manga, and to provide glosses (or sometimes even contrastive readings) for kanji. The Roman alphabet is also sometimes used to write Japanese. As Robertson (2015: 206) notes, "all four scripts carry various associations which can evoke certain feelings or images." Among the adjectives that native Japanese speakers associate with *katakana*, Robertson (ibid.: 207) lists the following: "novel, foreign, imitative, emphasizing, hard, simple, inorganic," as well as "futuristic" and "jarring."

Beichman's; but when we consider this shift together with the other features of her translation discussed above, the case for choosing hers as the ostensibly foreignising translation seems all the stronger.

As for which translation should occupy the **'ostensibly domesticating'** slot in the pair, I chose Rexroth and Atsumi's. Several factors led me to this decision. First, its line breaks all coincide with either a major constituent boundary or conventional punctuation, features which would be expected to aid readers in quickly or easily processing the text. As Venuti ([1995] 2008: 49) indicates: "Fluency assumes a theory of language as communication that, in practice, manifests itself as a stress on immediate intelligibility […]." Second, although Rexroth and Atsumi italicise the promise to eat the shellfish, in using 'eat' ('*I will eat them all when day breaks*') rather than 'gobble' or some other semantically near word, there does not seem to be as much of a shift in register as in Beichman's translation. The third reason is related to the second: compared to the other translations, Rexroth and Atsumi's generally seems to hew the closest to what Venuti ([1995] 2008: 4) understands to be a "standard" kind of English—that is, neither "colloquial ('slangy')" or "archaic." Thus, while Beichman's translation arguably tends more toward the colloquial ('dead of night,' 'couldn't help it had to [sic]'), and Morton's takes a slightly more formal tack ('In the night I awoke,' 'I passed the night in sleep'), Rexroth and Atsumi's seems to use language that is less marked in terms of register or formality.

The next translation I discuss is Elliott and Kawamura's (1998: 17-18) translation of Tanikawa Shuntarō's "Nijūoku kōnen-no kodoku" ([1952] 1967: 401). I categorised this translation as **ostensibly foreignising**. To begin with, I encountered two other translations of this poem: Bownas and Thwaite's (1964: 231) and Takako U. Lento's (2011: 47). I present these translations for reference below:

Two Billion Light-Years of Solitude

Human beings on this small orb

sleep, waken and work, and sometimes

wish for friends on Mars.

I've no notion

what Martians do on their small orb

(*neriri*ing or *kiruru*ing or *harara*ing).

But sometimes they like to have friends on Earth.

No doubt about that.

Universal gravitation is the power of solitudes

pulling each other.

Because the universe is distorted,

we all seek for one another.

Because the universe goes on expanding,

we are all uneasy.

With the chill of two billion light-years of solitude,

I suddenly sneezed.

(Elliott and Kawamura 1998: 17-18)

Alone in Two Billion Light Years

On this small sphere

humans sleep, wake, work

from time to time want friends on Mars

I don't know what Martians do

on their small sphere

(maybe they sleep'eep, wake'ake, work'ork)

but from time to time they want friends on Earth

that's absolutely for sure

Universal gravitation is

the force of being alone, attracting each other

The universe is warped

that is why all of us seek each other

The universe is growing fast

that is why all of us are uneasy

Standing alone in two billion light years

I sneezed, in spite of myself

(Lento 2011: 47)

The Isolation of Two Milliard Light Years

The human race, on its little ball,

Sleeps, wakes, and works,

Wishing at times for companionship with Mars.

The Martians, on *their* little ball--

What they do, I don't know.

Maybe they *sloop*, *wike*, and *wook*.

But at times they wish for companionship with Earth--

That's certain.

Universal gravitation

Is the pulling together of the force of isolation.

The universe expands

And so we all unite our wants.

The universe distends

And so we are all uneasy.

The isolation of two milliard light years

Prompts an involuntary sneeze.

(Bownas and Thwaite 1964: 231)

Deciding which translation was most consistent with Venuti's notion of foreignisation

was difficult, but Lento's emerged as an early contender. To employ Venuti's ([1995]

2008: 177) terminology, several features of this translation "frustrate[d] immediate

intelligibility" for me on my own first reading: its lack of line-final punctuation, coupled

with the inconsistent use of capitalisation to begin new sentences, forced me to backtrack several times to revise my understanding of how certain lines fit together grammatically. However, Elliott and Kawamura's translation and Bownas and Thwaite's translation seemed to manifest a more obvious contrast in their handling of the three 'Martian' words which appear in line 6 of the ST.

Tanikawa records these words in *katakana* and appends inflections of the verb *suru* ('do') to them in *hiragana*, as would conventionally be expected. By contrast, the entirety of lines 6-7 of Ishigaki's poem, containing the threat to devour the clams, is in *katakana*, verbal inflections included. In this way, Tanikawa's use of *katakana* arguably signals the otherness of the individual items themselves, whereas Ishigaki's use sets off the entirety of the lines in question. That is, Tanikawa's use of *katakana* discriminates between 'imported' and native Japanese elements, whereas Ishigaki's use makes the entire sentiment of her lines somehow 'other.' Coincidentally, Tanikawa's Martian words are 'imported' into Japanese as verbal nouns—these are nouns which take *suru* and can thus function as verbs. As Tsujimura (2014: 140) writes: "Many verbal nouns in Japanese come from Sino-Japanese compounds, but they also include loanwords […]." Some English loans which function as verbal nouns are *kisu* ('kiss') and *doraibu* ('drive') (ibid.). However, some loans are more "acclimatized" than others (Miller 1967: 246). As Miller (ibid.) indicates, "[m]ost completely acclimatized are, presumably, those Chinese loanwords that have ended up being inflected according to Japanese morphology." For instance, Heian-period (794-1185) texts "commonly treat *sōzoku* 'formal dress' in this way, with such forms as the gerund *sōzokite* appearing" (ibid.). Some loans of non-Chinese origin seem to have achieved this status, as well, e.g. *daburu* ('to be doubled, duplicated') from the English 'double'; rather than taking *suru* as a verbal noun, the loanword itself inflects, e.g., *dabutta* ('was doubled'). In the choice of

script and part-of-speech that Tanikawa assigns them, the Martian words are 'imported' in a way that signals their foreignness while at the same time maintains a certain phonemic correspondence with their native Japanese 'translations' in line 2 of the ST.

Elliott and Kawamura's and Bownas and Thwaite's translations take two different approaches to the Martian words. Although they are nonsense words made to stand for an imaginary language, they are *Japanese* nonsense words made to stand for an imaginary language. The only translation to highlight this fact is Elliott and Kawamura's, which transliterates them and affixes '-ing.' Moreover, just as Tanikawa's text uses heterogeneous script types (*katakana* and *hiragana*) for these words, so Elliott and Kawamura's translation uses a mix of italics and non-italics—within each orthographic word—such that the Japanese part of the words is italicised and the English part is not. Thus, within single orthographic words, we find a mix of two different glyph types, reflecting typographically the heterogeneous origins of the words. Bownas and Thwaite's translation, on the other hand, sticks to English phonotactics: 'Maybe [the Martians] *sloop*, *wike*, and *wook*.' In using items with phonemic correspondences to 'sleep,' 'wake' and 'work,' Bownas and Thwaite are arguably "doing as the first writer did, rather than repeating what he did," to borrow a phrase from Barbara Folkart (2007: 29). Just as (but of course not *exactly* as) Tanikawa presents the sounds of his own language as foreign, so Bownas and Thwaite set apart these (phonotactically) English words as loans.

'Doing as the first writer did' is a strategy that Folkart associates with "poetically viable" translation, as a superior alternative to the "foreignizing approach" (ibid.: xi). Indeed, Folkart advocates a "writerly" (ibid.: xii) approach to poetry translation, focused on the "proactive, 'making' forces that drive poems into being […]" (ibid.: 30). One of these 'forces' involves what she calls, after Eco (1975), "ratio difficilis"—or "what the

poet has done with her raw material to re-motivate it into a poetically viable sign"
(Folkart 2007: xii). Folkart's attitude is that if this results in a 'domesticating'
translation, so be it: indeed, she makes it clear that, for instance, Morel's translation of
Joyce's *Ulysses* "owes its viability" (ibid.: 214) to the domestications it performs (see
ibid.: 214ff). (Unfortunately, Folkart also tends to universalise, establishing *her*
responses to texts as a kind of universal measure of poetic viability.)

On the other hand, Bownas and Thwaite's approach to the Martian words can also
be read as 'minoritising' to the extent that it "open[s] up the standard dialect and literary
canons to what is foreign to themselves, to the substandard and the marginal" (Venuti
1998: 11). What I mean is: two of the words they use ('sloop' and 'wike') already exist
in English; and while 'wook' does not have an entry in the *OED*, it resembles an attested
historical spelling of 'week' ('wooke'). While I would never describe an item or
structure *per se* as 'substandard,' 'sloop' (a kind of small ship) could be conceivably
described as coming from a marginal discourse—the vocabulary of maritime transport.
'Wike,' on the other hand, meaning "a corner of the eye or mouth" (*OED*) is listed as
"[n]ow only *dialect.*" However, in the context of the translation itself, the sound-
similarity of the three 'Martian' words to the English words 'sleep,' 'wake' and 'work'
make their ostensible meaning *more* obvious, arguably making the text easier to
process—hence supporting the interpretation that this feature is domesticating.
Furthermore, Bownas and Thwaite employ italics in line 4 to mark information focus
('The Martians, on *their* little ball—'); the use of italics has been interpreted as a way to
"facilitate interpretation and increase fluency and idiomaticity" (Saldanha 2011b: 439).
Bownas and Thwaite's punctuation also resembles what Baron (2001: 23) calls the
"rhetorical tradition of punctuation": in this tradition, the purpose of punctuation is
"mainly to aid the reader in dividing up text for subsequent oral delivery, but also to

help clarify meaning" (ibid.: 23-24). For example: 'The human race, on its little ball, /

Sleeps, wakes, and works […]' could just as easily be 'The human race on its little ball /

Sleeps, wakes and works […].' However, the punctuation in line 1 arguably makes the

contrastive emphasis in line 4 ('The Martians, on *their* little ball—') that much more

obvious. To the extent intonation and punctuation work together to stabilise meaning,

this seems more recuperable to the category of domestication that foreignisation.

That said, one could argue that Bownas and Thwaite's use of the term 'milliard,' a

word that has since largely fallen out of usage in British English, satisfies the criterion of

"poetical archaisms" which Venuti ([1995] 2008: 121) associates with foreignisation. As

we can see from the Google Ngram chart below—based on data from the "British

English (2019)" corpus, with 'smoothing' set to the default value 3—the last decade in

which 'milliard' was used consistently more frequently than 'billion' was the 1920s.

Fig. 6. Google Ngram Chart: 'Milliard' vs. 'Billion'
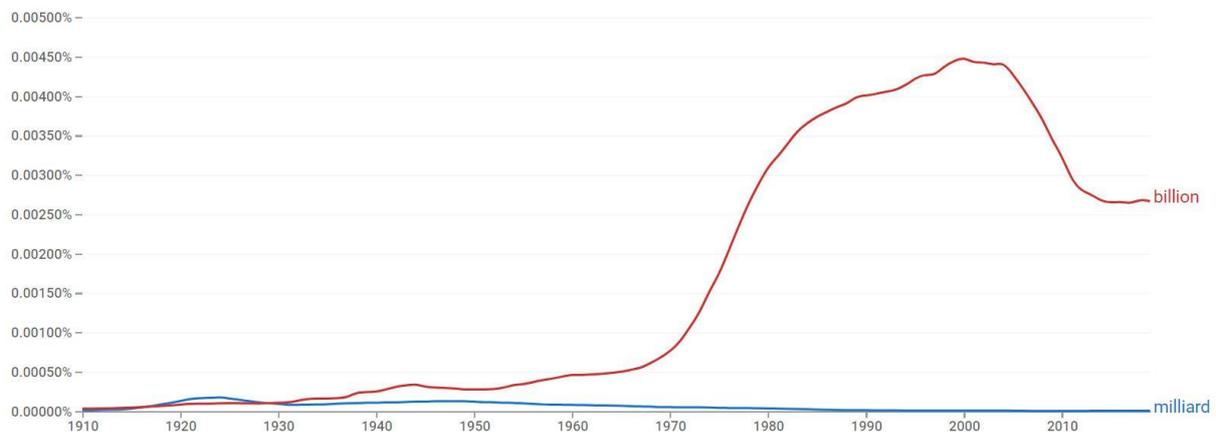
This calls to mind Tymoczko's (2000: 38) most salient criticisms of foreignisation: "[…]

how does the passage of time affect the quality of resistance? Once resistance, always

resistant? Or is resistance related to the specific historic and cultural moment of a

translation? […] Can we fault translations of the past for not being sufficiently

resistant?" Inasmuch as Venuti seems to definitively assign translations to one category or the other based on their original reception scenarios, one might call the use of 'milliard' a domesticating feature. However, in the more recent *New Edition* (1998) and *Revised Edition* (2009) of *The Penguin Book of Japanese Verse*, "milliard" has been replaced by "billion" (amid several other changes), suggesting an awareness that the word might disrupt readers' engagement with the text—hence supporting an interpretation of this feature as foreignising. This is a decided drawback to the current methodology: a more interesting question might be to ask whether younger respondents will tend to underline this feature more often than older respondents—insofar as the latter may be familiar with it from their youth. Alternatively, we might envision a scenario in which older respondents will underline it as a word that gives an impression of foreignness through its recent *disuse*. Then again, perhaps encountering the word in the title will activate for readers a different set of lexical probability expectations altogether, so that once readers accept that the text comes from an earlier time, the word no longer stands out. If this categorisation of Elliott and Kawamura's and Bownas and Thwaite's translations (as 'ostensibly foreignising' and 'ostensibly domesticating,' respectively) is not fully satisfactory, we hope that it may be allowed on the grounds that: 1) finding competing translations of modern Japanese poetry that differed even to the degree that these two texts did was difficult; and 2) this research is exploratory—we must make *some* predictions even if we are not totally confident in them.

The next text to discuss is Sato's (1981: 433) translation of tanka no. 145 from Yosano Akiko's *Midaregami* (1901). For reasons discussed below, I have classified this translation as **ostensibly foreignising**. As with the previous poems, I compared three competing translations. The other two were: Janine Beichman's (2002b: 214) and Reichhold and Kobayashi's (2014: 99). I present these translations below:

This evening rainfall is merciful. Traveler, my love, don't ask for a shortcut, but take a room with me

(Sato 1981b: 433)

dusk falls

with the blessings of rain

don't go

or ask for a shortcut

just find some lodging

(Reichhold and Kobayashi 2014: 99)

What falls tonight is the rain

of love's desire

Dear traveler

do not ask the shorter way

but make your lodging here

(Beichman 2002b: 214)

My reasons for describing Sato's translation as ostensibly foreignising have to do largely (though not exclusively) with his decision to translate tanka as a single line. This strategy, adopted by Sato for the first time in *From the Country of Eight Islands* (1980), proved controversial then and is still remarked upon today.

For example, as Arntzen (2008: 156) writes in her review of Sato's *Japanese Women Poets: An Anthology* (2007): "The one-line format takes a bit of getting used to as a reader. Sato asserts that he leans toward literal translation […]. Although he tries even to avoid punctuation, which was not present in the Japanese classical originals, he is forced to use commas and semi-colons to suggest the caesuras in the original syntax." Arntzen (ibid.: 156-157) goes on to compare Sato's translation of a tanka by Izumi Shikibu to a five-line translation of the same by Hirshfield and Aratani (1986), and concludes that "Hirshfield's translation is easier to understand at first glance, but Sato's captures the compressed and tense syntax of the original." From the standpoint of forestalling "immediate intelligibility" (Venuti [1995] 2008: 49), then, this would seem to support the classification of the one-line format as foreignising. Moreover, Arntzen's (2008: 157) conclusion that "[r]ead slowly, the one-line format can be powerful" suggests that such an approach *can* be aesthetically viable (though perhaps not consistently).

Non-professional reviews that discuss Sato's one-line technique are perhaps slightly more mixed. As one reviewer on the social book recommendation site *Goodreads* writes (*apropos* Sato's book *On Haiku* [2018]): "Sato brings this [one-line] approach over to English, ignoring line breaks and punctuation entirely (barring the occasional colon). Sometimes it works, but other times the effect feels needlessly janky. Poems that sound incredible in the original Japanese sometimes lose their original appeal in Sato's handling." Meanwhile, a reader of Sato's *String of Beads: Complete Poems of Princess Shikishi* (1993) writes in an *Amazon* review that the one-line approach to tanka "makes for an interesting experiment, and I approached the idea with an open mind as I started reading, but after finishing the book I have to say that in general it is not a very effective approach." They continue: "Tanka have no rhyme and

no meter, so the only thing making them poems really is their musical rhythm—and perhaps the gracefulness of their imagery. The one-line format in English though totally flattens out any potential rhythm so that more often than not the poems end up sounding like flat, declarative statements" (ibid.).

Statements like the above would seem to corroborate the intuition that English-language readers are more accustomed to multi-line translations of short Japanese poetic forms (particularly haiku). As J. Keith Vincent (2020: 108) has it in his review of *On Haiku*, "[b]y now, the three-line format has become so ingrained in English that most readers would be surprised to find it doesn't work that way in Japanese." I do not think that anyone would dispute the observation that haiku is more well-known among Anglophone readers than tanka. In this sense, William LaFleur (1983: 196) was probably correct to say—in his contemporaneous review of *From the Country of Eight Islands*—that "the common reader who picks up and reads this anthology will probably remain totally unaware that Sato's tanka in English are a break with [translation] convention" (though I appreciate that 'common reader' is a problematic term). However, what was interesting for LaFleur was not just the disturbance this represented to "the ordinarily quite quiet waters of our discipline" (ibid.), but also the difficulties such a strategy might represent to TL readers.

While contemporary reviewers critical of Sato's new one-line approach included Anthony Thwaite (co-translator of the Tanikawa poem discussed above) (1981), Phillip Harries (1982), Earl Miner (1982) and Karen Brazell (1983), it was William LaFleur (1983) who was probably the most outspoken—and his review is probably the most interesting to consider from a Translation Studies perspective. Specifically, LaFleur (ibid.: 199ff) offers what he believes are three "good and compelling reasons why we ought not follow Sato in rendering tanka as one-liners." The first is "the fact that a one-

line poem—at least in Western languages—is willy-nilly at the same time a no-line poem" (ibid.: 199). He continues:

> This is a fact around which there is, I think, no route of escape. Lines in poetry are in this respect like sexes in the world of biology: you need to have at least two to make the whole question of sex a meaningful one. Two would seem to be the lowest common denominator if we are going to speak of lines of verse at all; there are couplets and there are parallel lines, but to speak of a "one-line poem" is to speak of something that cannot exist.
>
> (LaFleur 1983: 199-200)

In positing the 'one-line poem' as a kind of unnatural category, LaFleur (ibid.: 200) betrays what would now be considered a wildly regressive attitude toward sex and gender: "The difference between one and two [lines] in this matter is immense. It is, ultimately, the difference between prose and poetry. One may detect poetic 'elements' in certain kinds of prose and one often has the impression that certain poems are terribly prosy—a situation not unlike the existence in our world of 'feminine' men and 'masculine' women." While I lack the space to unpack the implications of this gender view, it is worth pointing out that single-line haiku by the American poet Marlene Mountain (1939-2018) predate Sato's translations by four years (see Mountain's 1976 book *The Old Tin Roof*). My point is that, whether LaFleur likes it or not, single-line output had been both presented and received as poetry in Anglophone contexts before Sato—and it continued to be afterward. Indeed, as Kacian (2013: 347) notes, the major American poet John Ashbery "was impelled by Hiroaki Sato's one-line translations of haiku to try his own hand at the genre." Ashbery even subsequently provided the following blurb for the book: "*From the Country of Eight Islands* has been my pillow

book since it came out. It is one of the greatest books of poetry I know." This alone would seem to refute LaFleur's second criticism of one-line translation: "Margins on a page may make it possible to detect a poem there, but they do not themselves make for the kind of 'heightened speech' which poetry must be if it is to be something worthy of aesthetic attention, that activity which we call 'appreciation', and retention as something of cultural value" (LaFleur 1983: 201). It is notable how LaFleur pivots from a formal to a qualitative definition of poetry. His first argument declares Sato's one-line translations *not* to be poems (because one-line poems do not exist in English), while his second argument allows that they *might* be poems but states that they certainly do not rise to the quality of *good* poetry worth retaining.

However, what is interesting about LaFleur's case from a TS perspective is how *intentio lectoris* and expertise interact to reinforce a target-oriented view of translation. The shift from a lower-case to an upper-case (so to speak) conception of poetry seems premised upon a willingness to universalise one's own experience of a text or text-type, establishing that experience as a universally valid basis from which to argue about how *all* texts of that type *should* be. (Expertise seems to empower *intentio lectoris*.) In this case, the text-type is Japanese tanka translated into English, and the supposedly ideal form this should take is five lines. This is because one-line poems do not exist in English (or else are inferior poetry, according to LaFleur's expert opinion). LaFleur's expertise allows him to affirm that this is an essential difference between Japanese and English-language poetry. Because he has not encountered a one-line poem in English that 'works' for him—that rises to the level of "something worthy of aesthetic attention" (LaFleur 1983: 201)—he takes lineation to be a *sine qua non* for any English-language poem that deserves the name poetry. Now, LaFleur's expertise in the SL would seem to stand him in good stead with regard to SL *intentio auctoris*: indeed, he surmises that

"the non-lineated presentation" of tanka in Japanese "may have been [the] actual practice" of tanka poets, but "it was not an *artistic* norm of their practice as poets" (ibid.: 205). "Thus," he continues "there is no need to hold that it must become one of our norms when we render their verses into our own modern languages" (ibid.). This is LaFleur's third argument against Sato. As we can see, it shifts the emphasis from LaFleur's expertise as a reader of TL poetry (which allows him to universalise his experiences of English-language poetry), to his expertise as a reader of SL poetry (which allows him to appeal to the notion of *intentio auctoris*). However, in appealing to *intentio auctoris* as a constraining force on *intentio translatoris*, he inadvertently forecloses the very potentialities he claims to want translators to bring to the TL: "That from the donor culture must, through the translator's art, enter as much as possible into the literary norms, modes, and potentialities of the recipient one" (ibid.: 205). In order to promote those kinds of change, it would seem necessary (at least according to the model presented here) to use one's SL expertise to insist on the freedom inherent in the translator both as a reader of the ST and as a writer of the TT. This is perhaps a counter-intuitive use of expertise, as it concedes that the norms or expectations to be contravened may be one's own.

Nevertheless, LaFleur is right about one thing: the *relative* scarcity of one-line poems in English. That is, if someone constructed a representative corpus of English-language poetry (either original or translated), we would probably expect to find many more multi-line poems than single-line ones. In this respect, Sato's translation seemed more likely than the 'competing' translations to challenge values in the receiving situation (i.e. in the responses to my surveys) about what constitutes a poem. I was somewhat buoyed when a colleague provided a reaction that seemed consistent with Venuti's description of foreignisation as "stag[in] an alien reading experience" ([1995]

2008: 16). Flipping through a copy of a survey, they did a double-take when they came to the page with Sato's translation. After I explained that Sato's translation reflected the conventional layout of tanka, they nodded and said: "I was thinking, where's the poem?"

But it is not just the challenge presented by the layout that influenced my categorisation. I was also conscious of the potential impact of the point of view of the poem. It has been remarked, though I cannot remember by whom, that Japanese love poetry in the waka or tanka tradition has tended to focus on the *speaker's* experience of longing or desire, while canonical love poetry in the West (an inexact designation, I know) has emphasised positive attributes of the *beloved*. As Waley (1965: 114) wrote: "Most Japanese love-poems deal with separation and its sorrows […]." Naff (1979: 44) summarises Brower and Miner's (1967) work on the "conventionalized course of love in the imperial anthologies" as follows:

> There is first an awakening awareness of the beloved, usually through rumor, very rarely through sight; then an exchange of letters […]. At this stage the man fears he will not be successful and the woman that the man may not be sincere or that she might become involved in a scandal. Next comes the consummation, which is followed next morning by an exchange of poems, as are subsequent secret meetings. After a time, the man's interest begins to cool and finally the woman is left to sorrow, resentment, or despair, states in which were written the most powerful and moving love poems.

Where canonical love poetry in English is concerned, Julieta Flores Jurado (2020: 36) argues that Petrarchism has been a "master discourse in the Western lyric." Flores Jurado discusses Petarchism as a mode in which it is typically a man who "voices

167

admiration" while the woman is "admired and silent" (ibid.). Yosano Akiko's poetry, on

the other hand, would seem to challenge both lyric paradigms—both the canonical

Japanese and Western traditions. As Beichman (2002a: 201) puts it, "[u]nlike most

classical Japanese love poetry, which wove its narrative around moments of longing, a

brief period of union, and the woman's abandonment, many of the love poems of

Tangled Hair are constructed around the peak of satisfaction: its forecast, its memory, or

its actuality." Out of consideration for my survey respondents, I have chosen a much less

risqué example than I might normally have, but which nonetheless could be described as

departing from whatever stereotypes may still linger in the public imagination about

Japanese women as unassertive or passive. Given that Beichman stresses the

"directness," "open self-assertion" (ibid.: 92) and "bold sensuality" (ibid.: 112) of

Akiko's poetic voice, topics which Beichman treats so perceptively in her account of

Midaregami, I was surprised to find that boldness slightly more in evidence in Sato's

translation of tanka no. 145 than in Beichman's own—though I should hasten to add this

is not in any way a comment on the value of either of their approaches. Whereas

Beichman renders the final two units of the tanka in almost courtly terms ('do not ask

the shorter way / but make your lodging here'), in Sato's translation the appeal becomes

more colloquial, more direct ('don't ask for a shortcut, but take a room with me'). This

marks a shift in register from earlier in the poem. Indeed, we might pinpoint the shift to

the middle of the poem, where there is a mismatch in the "degree of evaluation" (Quirk

et al. 1985: 774) between the two vocatives "Traveler" and "my love"; that is, one is

more intimate than the other. In this sense, the discourse used is Sato's translation is

more heterogeneous than Beichman's, which is another argument for categorising it as

'ostensibly foreignising.' Finally, I would be remiss not to mention the American

spelling of 'traveler' used in both Sato's and Beichman's translations: I have not altered

it. Indeed, Venuti describes the "recasting" ([1995] 2008: 157) of American translations into British English for British audiences as an example of "assimilat[ing] the foreign text to dominant linguistic values […]." Given that I seek data from native speakers of British English, we would expect this item to give an impression of foreignness in the underlining task for Survey 2.

This left me with the difficult choice of whether to include Beichman's or Reichhold and Kobayashi's translation as the 'ostensibly domesticating' one. Both are presented as five-line poems; in this respect, neither really departs from conventional approaches to the translation of tanka—although it should be said that Beichman does not always use the five-line format, and has indeed convincingly argued for alternative lineations and visual arrangements (see e.g. Baichiman [2017: 73-75] in Japanese). However, Beichman's translation, in its use of capital letters, arguably gives more clues to its intonation contours than the rather choppier lineation of Reichhold and Kobayashi's. In other words, it was easier for me to tell where discrete utterances began in Beichman's text. In contrast, on first reading Reichhold and Kobayashi's translation, it was difficult for me to guess which lines would spill over into the next. In light of this, I had to conclude that Beichman's posed fewer challenges to readability, and thus should occupy the domesticating spot.

The final poem I have chosen for my survey is "Shunkan" by Yoshihara Sachiko. Here I discuss my translation of the text, and how I posited survey questions related to foreignisation. For the sake of reference, here is my translation, followed by Miller and Kudo's (1987: 57):

instant
瞬間
shunkan

The sea is terminal,
　　　日が The day is a casualty.
　　　　　　暮れる Here's dusk.
The moon is a fatality,
　　　夜が The day is terminal.
　　　　　　明ける Here's dawn.
Time is a fatality.
　　　人も The day goes dead.
　　　　　　死ね And so should we.
Running out—
　　　時が But always coming back—
　　　　　　またくる Time is generous with itself.
If the majesty
　　　うつ Of the moon and sea
　　　　　　くしさ Is saved across death—
No less vivid,
　　　のこる No more somber
　　　　　　なら Than before—
Then it must be
　　　hito-mo
　　　人 も The ugly few are we,
　　　　　shi-ne
　　　　　死ね And death is what we need.

(Garza)

Instant

The sea dies

Today dies too

The sun sets

The moon dies

Today dies too

The night lifts

Time dies

Today dies

People die too

170

Remorselessly

Dying repeatedly

Time comes again

If the beauty

Of the sea that dies

Of the moon that dies

Remains

Unfaded

Undarkened

Then people alone

Are ugly

Drop dead, people!

(Miller and Kudo 1987: 57)

Motivated by a desire to test the limits of the aesthetic viability of strict adherence to ST features—and partially inspired by the experimental translations of Sawako Nakayasu (2011) and Pan and Rock (2016)—I included in my translation some strings of text from the ST, in Japanese script. Both the choice of text and specific translation strategies were informed by Venuti's theory of foreignisation, as I describe in more detail below. There were several other factors informing my translation. First, this was my first time translating specifically toward psychometric scales. That said, perhaps this is not so remarkable when we consider that norms and constraints of various kinds may be a factor in any translation work. So perhaps what was novel at this stage of my experience in literary translation was rather the explicit-ness of those constraints, in the form of the specific Likert items.

In positing survey items for Survey 2, I was conscious of two notions discussed by Saris and Gallhofer (2007: 15-18): concepts-by-intuition and concepts-by-postulation. As examples of the former, Saris and Gallhofer (ibid.: 15) cite "judgments, feelings, evaluations, norms, and behaviors." Meanwhile, examples of the latter "might include 'ethnocentrism,' different forms of 'racism,' and 'attitudes toward different objects'" (ibid.: 16). In short, "concepts-by-intuition are simple concepts whose meaning is immediately obvious while concepts-by-postulation are less obvious concepts that require explicit definitions" (ibid.: 15). The authors argue that "[c]oncepts-by-postulation cannot be operationalized directly in survey questions," but are "normally defined by some combination of concepts-by-intuition" (ibid.: 28).

It is probably clear by now that foreignisation would be a concept-by-postulation. As Venuti ([1995] 2008: 19) himself writes, "The terms 'domestication' and 'foreignization' indicate fundamentally *ethical* attitudes towards a foreign text and culture […]." Thus, the next challenge in designing a survey around foreignization lies in identifying the relevant concepts-by-intuition and changing these into survey items "indicating the requested concept" (Saris and Gallhofer 2007: 13). Thankfully, Venuti himself provides several concepts-by-intuition that could be used for this purpose. In the end, I focussed on three indicators of foreignisation as an effect related to the cognitive processing of readers: whether the translation "register[s] the linguistic and cultural differences of the foreign text" (Venuti [1995] 2008: 15), whether the translation departs from "transparent discourse" (ibid.: 13), and whether it challenges "receiving cultural values" (ibid.: 96). I have focussed on cognitive effects for two reasons: 1) in Venuti's model, cognitive effects are prior to other effects— readers must first perceive the "linguistic and cultural differences" [ibid.: 34] of a translated text before it can "create a cumulative ethical effect" [ibid.: 268]. Thus it makes sense to first investigate whether the cognitive effects claimed for foreignisation are actually produced. 2) Investigating the ethical effects of foreignisation arguably calls for a longitudinal design,

where the same respondents are evaluated both before and after their exposure to the stimulus texts to determine any change in attitudes. Given the difficulties I anticipated in enlisting respondents, I decided to try to maximise the research benefits of a one-off survey. I followed Saris and Gallhofer's (2007: 13) "three-step approach" to transform the above concepts-by-intuition into the following questions:

1) Does this poem feel like it comes from a non-English-speaking culture?

2) Does this poem read like it was originally written in English?

3) Does this poem conform to your image of typical Japanese literature or culture?

Many other questions were possible, but I was conscious of the need to keep my surveys short and approachable; indeed, in Survey 2 I would already need to retain the five questions from Survey 1, so space was at a premium. Nevertheless, the concepts-by-intuition that I chose seemed central to Venuti's construction of foreignisation.

My notebooks from the time record that I was aiming for a particular configuration of scores on the Likert scales when it came to my translation: the goal was to "score highly on ratings for impression of foreignness among survey respondents, *while simultaneously* scoring highly on [the] other evaluative measures." As I recorded in my notebook: "A tall order? For sure." The sense of trepidation was compounded by doubts about the feasibility and seemliness of the task: "What if the desired effect can be accomplished through no particular special effort on my part? Is it not a bit cheap to take credit [for any attested effects of foreignisation] only after the fact?" As an early-career literary translator, I was also worried about taking too many risks, i.e. deviating from too many norms. I had previously published translations of Yoshihara Sachiko in the online journals *Lunch Ticket* and *Asymptote* but had yet to place highly in any poetry translation contests, as I would later do in

2018 and 2019.[29] Still, my name was going to be attached to this translation, and *I* wanted to

be satisfied by it, even if many other people were not. I hoped that whatever I produced for

this project would not damage whatever small reputation I had built up. I also hoped that the

resulting translation would not hurt my relationship with the copyright holder, a relation of

the poet's. But it helped that I had something of a model to follow. The first time I read

Sawako Nakayasu's "translations" and "anti-translations" of the Japanese modernist poet

Sagawa Chika in a slim volume entitled "Mouth: Eats Color" (2011), I felt a strange frisson

which, as I wrote in my notebooks, seemed "consistent [with] Venuti's description of

foreignization, but which I'm not sure I would have described that way had I *not known* of his

theory." What was perhaps more remarkable was that I felt this way *despite* my familiarity

with the Japanese that Nakayasu frequently employs: the texts seemed to unfold in a place

normally off-limits, a place where certain boundaries (e.g. between ST writer and TT

translator, between languages themselves) had dissolved—where English was liable to turn

into French or Japanese mid-poem (see "PROMENADE" on page 11), and even the

vicissitudes of Japanese character-input were implicated in a kind of *henkan misu* (conversion

error) poetics (see "PROMENADE (P ろめなで)" and "PROMENADE (露命撫で)" on

pages 47 and 49, respectively). It was also a place where the writer-translator might playfully

cast doubt on whether what *looked like* a Chinese translation really was in Chinese (Nakayasu

2011: 86). This was perhaps what contributed the most to my sense of the 'foreign' in the

text: there was a peculiar sense of dislocation *amid* linguistic particularity. As Johnstone

(2011: 203) notes, "[b]e it nation, region, county, city, neighborhood, or block, place has long

been adduced as a key correlate of linguistic variation." I think this is where the feeling of

dislocation came into play: there was a sense in which the texts refused to be tied down to a

---

[29] In 2018, my translation of a prose poem by Kasuya Eiichi was commended in the Stephen Spender Prize. In 2019, I won first prize in the open category with a translation of a poem by Itō Shizuo. I have also since published in *Modern Poetry in Translation*.

particular place. However, it is not difficult to imagine that code-switching might also invest places (e.g., bi- or multilingual households, or geographical areas with a history of language contact) with particular meanings. As Johnstone (ibid.: 211) writes, "spaces become human places partly through talk, and the meanings of places shape how people talk." So it was not just a sense of dislocation I felt, but a sense of *creation of place* as well. I cannot pinpoint a precise through-the-looking-glass moment, but nevertheless I found myself *on the other side* (at least temporarily), in a kind of imagined place that was both utopian and, in a strange way, aspirational: utopian in the sense that it seemed to wrest the power of language away from purist ideologies of the nation-state,[30] and aspirational in the sense that I wished I possessed the linguistic competencies to fully belong there.

All of this informed my sense at the time of being torn between two impulses—of having to "navigate [between] the Scylla and Charybdis of Venuti and Folkart," as I noted to myself. I had recently read Folkart's *Second Finding* (2007) and found myself sympathetic to many of the author's points. Folkart admits that "the visibility of the translated poem […] can arise from a number of radically different translation strategies," among which she includes "the sort of grainy foreignizing translation that fixates on the microstructures of the source language"; but she writes that "[n]ot all of these strategies lead to viable outcomes—if one defines viability as the extent to which the translation can function as a free-standing poem, a text […] that really *is* a text" (Folkart 2007: 284). The gist of Folkart's argument against foreignisation is that "the focus seems most often to be less on the esthetically relevant structures that constitute the essential, idiosyncratic otherness of the source *poem* than on the *grain* of the source *language*—its linguistic microstructures—perceived, and valued (naïvely, I feel) as a carrier of cultural otherness" (ibid.: xi). In other words, "[w]hat the foreignizing approach fails to recognize is that poems are always written from more or less outside the

---

[30] See also Baynham and Lee (2019: 9) on "translanguaging."

language that constitutes their raw material, that they are inherently foreign, even to the language out of which they have been crafted" (ibid.: xi-xii). What is at stake for Folkart is the recognition of translation as writing. But at a deeper level it is the creation of "a model of *making*, and a model of *being-through-making*" (ibid.: xiv), in which a central part is played by what we might call *renewal through aesthesis*—the transformation of both language and self through the act of writing. But it is not just the writer-translator who is potentially transformed; when translations are "authentic" (ibid.: 129), they can transform readers, too. As she puts it: "what good to the target-language reader is a poem with the wit and music leached out of it by a pedestrian and repetitive approach?" (ibid.: 22). And later: "Poetic creation makes ruthless demands, in the name of authenticity. The already-said is never good enough […]. Ezra Pound was right: artists are indeed in the business of 'making it new'—not for the sake of newness, but for the sake of stirring us to feeling and insight" (ibid.: 42). Folkart's (ibid.: 3) notion of 'the already-said' is closely linked to linguistic norms:

> Where poetic discourse is inaugural, lignification (to borrow Jean-Claude Michéa's neologism) is the process through which living, breathing language necrotizes, hardening into fossilized remnants of itself. […] This process of fossilization is an inevitable stage in the life of language. Innovative uses of language invariably get standardized and resorbed into the already-said: the novel becomes the expected then degenerates into cliché. […] The already-said—or norm or idiom—is the social, sandwiched in between the universal and the individual […].

In this recourse to the vocabulary of norms and desensitisation, we find ourselves back in the realm of Russian Formalism, and its salvos against habitualisation. However, if Folkart's notion seems uncomfortably close to the mechanistic model discussed earlier, she at least restores a sense of agency to the "crafting faber" (ibid.: 441n13), be they ST writer or TT

translator. And she pays more attention than most to the writer-translator as someone whose instincts and intuitions are shaped by their experiences as a *reader*. As she puts it, "[w]riting is driven by intuition—the intuition of a competent poet being a more complex, more complete, more highly organized and finely tuned grasp of what makes a poem than anything a theorist can aspire to formalize" (ibid.: 13). Thus, the "translator who is competent to make a derived poem can safely prefer her own instincts to any other reading, I believe" (ibid.).

My background reading in Yoshihara was also part of my cognitive context when I translated the poem. Proceeding from the understanding that a translation is an interpretation (a point on which both Venuti and Folkart agree), I had revisited some secondary sources collected during an early stage of this project. The idea was that I might encounter some material that would help me *better* interpret my selected text. Despite my scepticism of the necessity of expert readings in producing foreignising effects, and contrary to Folkart's (2007: 84) admonition against "irrelevant erudition," the instinct to want to *better* interpret the text—to produce an acceptable interpretation—felt totally natural. In this respect, I am probably a product of the university system. That is, I feel like I still want to produce interpretations that will earn the approval of my colleagues. In my defence, the motivation here was not to establish my interpretation as the best or only one—rather, it was to find material about the poet and her works that might spark off or inspire more aesthetically viable translation choices. In particular, there were three publications: one was an essay by Yoshihara ([1975] 1983: 189) in which she claimed to have "no poetics," one was an article about the author which touched upon Yoshihara's treatment of time (Kuninaka 2000), and one was an essay by poet Itō Hiromi ([1981] 2012) on Yoshihara's use of *kyūkana* —the historical *kana* spellings and characters in common usage until the government's adoption and promulgation of a new set of orthographic rules in 1946.[31]

---

[31] For more, see Seeley (1991).

In the first publication, an essay entitled "Tōkai-ni tsuite" [on concealing oneself], the poet announces in the first line: "I have no poetics [*shiron*]." This was a rather startling claim to encounter, perhaps because I had already read so many essays about the distinctive features of her poetic output. Indeed, the statement might recall Barthes's ([1953] 1967: 77) concept of "zero degree" writing—a kind of "neutral writing" whose "aim […] is to go beyond Literature by entrusting one's fate to a sort of basic speech, equally far from living languages and from literary language proper," a kind of "transparent form of speech" which seeks "an ideal absence of style." Of course, as Wales (2011: 442) points out: "It is questionable, however, whether a work can have no style: the very 'absence' of a marked style can itself be seen to be stylistically significant." However, it soon becomes clear that in Yoshihara's case, she is using the term 'poetics' in the sense of a conscious "poetic method" (*shi-hō*): "In truth, I'm quite frightened of the prospect of being made to disclose [*hakujō*, also 'confess'] a poetic method […] that does not even exist" (Yoshihara [1975] 1983: 190). As a structuring principle, such a method belongs, she seems to suggest, to some part of herself which is involved in the writing yet with which she is unacquainted: "Inexpert [*fubenkyō*] as I am, I read with great enthusiasm the essays that others have written [on my poetics]. For one thing, I have a *personal* interest in discovering my 'unknown self' [*watashi-no shiranai watashi*]. I have only written the poems, and there is no way for me to know what I am doing (or have done) in them without other people pointing it out […]" (ibid.: 189).

Later on in the essay, she gives two important insights into her process. First: "Although I said I have no poetics, it also strikes me as a basic fact that a poem is something which comes about through its relationship not only with 'the writer' but 'the reader' as well. In order to leave the possibility of diverse interpretations [*yomitori-kata-no tayō-na kanōsei*] open, [it is necessary] to not fix every image, to allow [the words] 'I' and 'you' to stand for many other 'I's and 'you's—or any […]" (ibid.: 194). The second insight involves the

178

revision process. After drafting a poem or part of a poem, she "set[s] it aside" for anywhere between "several days and several months" (ibid.). It helps if "enough time has passed for me to have forgotten what I had jotted down" (ibid.). She continues: "I read it over as if it were someone else's words. There [will be] parts I am embarrassed about. […] There [will be] spots that are [too] direct, places where Yoshihara Sachiko is *too present*. I cross those out. […] I only keep what I do not resist as a reader" (ibid.: 194-195). Finally, she adds that as she repeats this process, she sometimes gets lucky and "someone who is not Yoshihara Sachiko" comes along with just the line she was looking for (ibid.: 195). And like this, the poem is eventually finished. While she makes it clear that her writing self is strongly informed by her reading self, it is interesting that when it comes to interpreting her own work, she de-emphasises her own *intentio lectoris*. In other words, *auctoris* is for her a temporary position she may fill—and she only feels she has performed this role properly when her reading self and her unconscious (writing) self collaborate to generate lines she deems satisfactory. Nevertheless, what struck me at the time was a feeling of identification with the poet which I could not help: I was surprised that her process so closely resembled my own for creative writing. Venuti ([1995] 2008: 237ff) may argue against *simpatico* in translation, but what this experience showed me was how hard it is to *avoid* identification—not *in spite of* pursuing a foreignising translation, but *because* of it.

In other words, there was an inkling at this point that I was looking for *permission* in the author's own poetics to justify my translation strategies. In fact, Rosemary Arrojo (1997: 28) had already levelled a similar criticism at Venuti's translations of Milo De Angelis: "De Angelis is not, in any way, being perversely abused as a poet but is, definitely, one of the organizing principles that directs and inspires the translator's options."

Nevertheless, like Venuti's De Angelis translations, I thought Yoshihara's poem would probably not be "a congenial poem" to bring into the TL (Venuti [1995] 2008: 251), seeing as

it is not a very 'pleasing' poem, but is rather death-obsessed. After I translated it, it struck me that I used what may be described as a minor discourse: the clinical-apocalyptic vocabulary of J.G. Ballard (in my emphasis on "terminal" this and that). I staggered the lines and inserted Japanese to slow down reading, in a sort of compensation for Yoshihara's use of *kyūkana*: Itō ([1981] 2012: 6) and others have remarked on the alternative pronunciations suggested by Yoshihara's historical *kana* usage, as well as the "strangeness" (fushigi-sa) (ibid: 4) of this technique, and how it seems to mark her work out as removed from contemporary poetry. Also the line staggering came from my own aesthetic practice—a short poem I had written in that format had recently won a local competition, and this form felt suddenly 'viable' to me, worthy of further exploration:

Fancy Dress on the Otley Run

The girl on the bus is captivated.
　　　"They're having fun," her father says.
　　　　　The gladiators sneak around a hedge.

Kuninaka (2000: 431) has remarked in the context of another Yoshihara poem— "Mujun" (contradiction)—that the "multi-layered perception of extremes" is an element of "both her view of the universe as well as [her] poetic method." The observation that she "grasps an instant as if it were eternal" (ibid.) seemed apropos when it came to translating "Shunkan" [instant], and I wondered how I might signal the sense of eternal return that was part of my interpretation of the poem. It strikes me now that my pattern of line staggering— returning the first line of each stanza to the left margin—may have been one way of doing that. Finally, I also used overt rhyme as another way to compensate for the anachronistic look of Yoshihara's orthography. Venuti ([1995] 2008: 5) quotes the poet Charles Bernstein

(1986: 225) on the "elimination of overt rhyme & alliteration" in "current middle of the road poetry […]." Since these are "trends" Venuti associates with "fluency" ([1995] 2008: 6)—and since he so frequently associates fluency with domestication—I thought this might be another feature that could be described as foreignising according to his theory.

<p style="text-align:center">***</p>

As the analyses included in this chapter demonstrate, there is an unavoidably subjective element to the designation of the texts as ostensibly foreignising or domesticating. In order to predict features that might contribute to a foreignising effect, one must have some pre-existing image or understanding of the population of respondents to be sampled—and be conscious of the degree of similarity (or lack thereof) between oneself and these respondents. Thus, it has been difficult to base the above categorisations on textual features. However, I hope it will be allowed on the grounds that the research is exploratory. Indeed, as I discussed earlier, there are aspects in which the translations used in my surveys could *all* be described as foreignising to some degree. The point is rather to begin to isolate textual features from the set Venuti associates with foreignisation.

In this section, I have discussed the basis for my categorisation of stimulus texts as 'ostensibly foreignising' or 'ostensibly domesticating.' I have also explained how I operationalized the concept-by-postulation known as foreignisation, and how I produced my translation of the Yoshihara Sachiko poem "Shunkan." In the next section, I explain my general method for identifying FG in the stimulus texts.

3.2. GENERAL METHOD OF STYLISTIC ANALYSIS

In the interests of comparability, the general procedure for my stylistic analyses follows Van Peer (1980 & 1986)—his Ph.D. thesis and its subsequent book version. However, I have innovated several additional procedures to take advantage of corpus-linguistic tools not available when Van Peer undertook his study. Like Van Peer, I identify FG features at three different linguistic levels, tackling the level of phonology first, followed by syntax and then semantics. As I mentioned in the introduction, four types of FG occur across the above three levels in Van Peer's model:

1) statistical deviation

2) determinate deviation

3) internal deviation

4) parallelism

First is 'statistical deviation.' As Van Peer (1980: 68) writes: "one often encounters in poetry a kind of deviation that is the result of 'rarity.' In such cases the string is not really deviant but only highly unexpected." For instance, given the incomplete phrase 'from beginning to …,' most readers would, he surmises, supply the word 'end' (ibid.). But what about the phrase "from beginning to beginning"? He explains: "It is clear that [this] does not violate any grammatical or other rule. But the word 'beginning' at the end is so rare that it will create surprise. In other words, by using an element in a context where its probability of occurrence is almost equal to nil, our attention is drawn to this unexpected element" (ibid.). Finally, he emphasises the determining role of context in this kind of deviation: "The 'rare' element may go completely unnoticed in another context" (ibid.).

Second is 'determinate deviation.' This is "the kind of deviation that results from the breaking of a rule, e.g. the violation of a linguistic rule, or the clear infringement of an

established cultural code" (Van Peer 1986: 18). Moreover, "[s]uch rules and conventions can be of a linguistic, literary, social, cultural, or other kind, but in all cases it must be possible to account for the departure in terms of an infringement of a rule that can be made explicit" (Van Peer 1980: 80).

Third is 'internal deviation.' This is a "departure from a norm that is set up within the boundaries of the poetic text" itself (ibid.: 67). He goes on: "As such internal deviation occurs against the background of the remainder of the text itself. This is possible because a poem may set up its own internal regularities, for instance rhyme-scheme, stanzaic structure, regularity in sentence-pattern, etc" (ibid.).

Fourth is 'parallelism.' This refers to "a pattern of equivalences and/or contrasts that are superimposed on the normal patterns of language organization" (Van Peer 1986: 23). He continues: "While deviance is the result of a choice the poet has made outside the permitted range of potential selections, parallelism is the opposite process, in which the author has repeatedly made the same, or similar, choices where the normal flux of language would tend to variation in selection" (ibid.: 23).

Of course, as Van Peer himself is aware, these categories are not unproblematic in themselves. One of the biggest unresolved questions regarding these categories is one of degree: for instance, where statistical deviation is concerned, exactly *how* deviant does a feature have to be to fall into this category? And by comparison with what extra-textual context? Furthermore, how 'similar' do two textual features have to be in order to count as an instance of parallelism? Moreover, where determinate deviation is concerned, is it not the case that 'literary,' 'social' or 'cultural' 'rules' can also be described in terms of normative expectations? And if this is the case, then how is this category different from statistical deviation? All of this points to an inescapable element of subjectivity in the process of analysis. To circumscribe the role of subjectivity as much as possible in my own analyses, I

decided at the start that in cases of uncertainty over categorization, I would ask myself what Van Peer had done in a similar situation in his own analyses. If he could obtain such strong results partly on the basis of his own subjective decisions, then in the interest of comparability, it made sense to follow his lead in matters of categorisation.

At the level of phonology, I first present a broad phonetic transcription. Then I tackle orthography and punctuation (insofar as these are related to the phonology of the poem). Next, I perform an intonation analysis, and examine how tone units and lineation interact. This is followed by a rhymico-metrical analysis. The next stage of analysis concerns word length—I compare the frequency and distribution of words of different lengths against the poetry subcorpus of the British National Corpus (BNC). Then comes an analysis of individual phonemes, as well as assonance, consonance, alliteration and rhyme. Finally, I examine consonant clusters and phonotactics.

It is important to address linguistic variety here. My transcriptions are in Received Pronunciation. I chose RP for a number of reasons: 1) Van Peer (1986) obtained his results in another Northern city (Lancaster) despite having transcribed his texts in RP, and despite two of the poems being by American poets; 2) I had no way of knowing beforehand which variety of British English respondents used; 3) even if I knew which variety a particular respondent used, it would be unfeasible to produce different transcriptions for different readers; and 4) as Roach (2004: 239) writes, this accent "has been used as the standard in phoneticians' description of the pronunciation of British English for centuries."

The most obvious objection to transcribing in RP is that, as (Knowles 1987: 4) writes, "relatively few people actually speak it." While this point is sensible, we must also recognise that speech varieties exist along continua. When we say that few people speak it, what we are really saying is few people speak an officially recognised version of it. This is what Trudgill (2002: 174) gets at when he remarks, "it only takes one non-RP feature for a speaker not to be

a speaker of RP." Moreover, RP's "pre-eminent status in broadcasting" (Roach 2004: 239) suggests that, while it is a minority accent, it is a rather influential one. Thus, Knowles (1987: 4) writes: "The vast majority of English speakers today have a standardized variety of English. In England their pronunciation is likely to be influenced by RP, but retains some local flavour." It is my position that we can recognise RP as a minority accent while still acknowledging that for the purposes of Van Peer's (1986) study, RP transcription was *good enough*. **Inasmuch as readers' responses largely conformed to his predictions— predictions he made on the basis of RP—it appears that influences of dialect or accent were minimal. This is especially notable given that three of the four stimulus texts Van Peer (1986) used for the strikingness portion of his study used overt end-rhyme.**

At the level of syntax, I first present syntax trees of every sentence in the text. I use these to check for anything that might enter our FG inventory, with an eye toward features like ellipsis, fronting, structural ambiguity and parallelism in constituent structure. Next I discuss tense, aspect and mood. This is followed by an examination of lineation and punctuation as they relate to syntax. Then I discuss syntactic complexity, and patterns of parataxis and hypotaxis. Finally, I examine the 'keyness' of individual parts of speech with reference to the poetry subcorpus of the BNC. (For an explanation of 'keyness' see section 4.1.3.)

At the level of semantics, I first perform a keyness analysis of lexis, using the BNC as a reference corpus and adjusting for the so-called 'aboutness' of the text, as well as frequent collocates. Both of these concepts are discussed in further detail in section 4.1.3. Next, I look for instances of FG related to conceptual metaphors, schema violations, implicature, register, anaphora, cataphora, logical contradictions, and other such issues. The rationale for these procedures, as well as in-depth discussion of how they were performed, can be found in the sample analysis in Chapter 4. As in Van Peer (1986), analysis at each level was performed separately. Moreover, I finished these stylistic analyses before analysing (or even looking at)

any completed surveys.

3.3. DESCRIPTION OF SAMPLE AND SAMPLING METHOD

There were two populations of interest: 1) adult native speakers of British English[32] living in the geographical area covered by the 2017/2018 BT phone directory for Bradford, and 2) undergraduate native speakers of British English enrolled in literature courses at the University of Leeds. This choice was a multi-faceted one, informed by both methodological and practical considerations. Below I explain the reasons for these choices, the required vs. actual number of respondents, the implications of my sampling methods, and the distribution of respondents in each survey.

The choice of a population is crucial as it relates to issues of replication and progress in social science. In a direct replication, a new researcher "seeks to duplicate the sampling and experimental procedures of the original research," whereas in conceptual replication "the original methods are not copied but rather purposefully altered to test the rigor of the underlying hypothesis" (Makel et al. 2012: 538). Since Van Peer (1980 & 1986) studied the responses of undergraduates, my study also needed to include undergraduate respondents for the sake of comparability. However, one motivation of my study was to see whether diversity of response would emerge if we opened participation up to respondents *beyond* undergraduates—in this sense, my study always demanded a conceptual replication.

---

[32] See also section 3.2. I chose this criterion to help control the potential influence of language variety on survey results. It was also the principal screening question in the survey. Some respondents were also screened out on the basis of their written comments: these sometimes indicated that respondents did not answer the questions as posed. Several other reader variables were accounted for by a questionnaire at the end of the surveys: age, education level, familiarity with poets or poems, number of hours per week spent reading for pleasure, level of literary training, familiarity with the theory of FG, familiarity with foreignisation, and attitudes toward poetry.

There will obviously be differences between the student population I sampled and the one Van Peer sampled at the University of Lancaster in the late 1970s; but even if I sampled undergraduates at the University of Lancaster *today*, we would expect to find important differences between the current and past populations. There were also practical considerations: since I lived closer to the University of Leeds, it was more time- and cost-effective to sample students there (no small consideration, as I bore the cost of surveys myself). My location also helped preserve one dimension of comparability between my student respondents and Van Peer's—they were all students attending university in a Northern city.

Sampling from a Northern area offered an important theoretical affordance, particularly when it came to the foreignisation surveys. This involved the intuition that familiarity with Japanese culture might be lower in the North than in the South, given the headquartering in London of most of the major organisations related to UK-Japanese relations and cultural exchange. This intuition was partially confirmed when I heard a representative from the Japan Foundation address "regional differences" in the British awareness of Japan: in a BBC Front Row radio interview, Junko Takekawa relates her concern that people outside of London may remain at a "very basic level of the knowledge of Japanese culture" (BBC Front Row 2019). Further support comes from Kate Taylor-Jones, one of the organisers of Japan Now North, a festival first launched in Sheffield in 2018. Reporting on the 2019 festival, she writes: "Sheffield still suffers from the same fate that many of the Northern cities do—the continual and never-ending focus on London as the nexus around which much cultural activity in the United Kingdom is based. This is particularly an issue when it comes to the cultural output from Asia" (Taylor-Jones 2019: 203). If familiarity with Japan is indeed lower in the North, then we might expect the effects of foreignisation to be more easily observable in a

sample of Northern respondents.

As in Van Peer's (1980) study, canvassing of student respondents was "done through administrative channels, and contact with the informants prior to the experiment did not occur" (112-113). Surveys were collected in this fashion from three sources: a level-one World Literature class, a level-two Chinese literature class, and a level-three World Literature class. Participants were initially instructed to complete the surveys at home, but this resulted in a very low response rate, so instructors offered to administer the surveys in class. I provided each instructor with the same script for introducing the surveys. Since no feasible method of random sampling existed, and access to students depended on the amenability of instructors, this sample was a convenience sample. In this respect, the sample is also similar to Van Peer's (1980). However, Van Peer (1980) was also able to enlist undergraduates majoring in subjects other than literature. As I was dissuaded at the ethical review stage from incentivising my surveys, and I had already spent considerable money printing and posting surveys to non-students (see below), I decided that my remaining resources were perhaps best focussed on the above groups of students.

Since I was sampling students in Leeds, representativeness demanded that I should also sample non-students there. The nearest phone directory I could obtain was for Bradford—its residential listings included parts of the metropolitan borough of Leeds. While a phonebook does not include every adult living in a specified area, this was the best available sampling frame, and it would allow me to make an effort at random sampling.

Using effect sizes derived from results reported in Van Peer (1980 & 1986), I performed power calculations based on Cohen (1988) to determine sample sizes required to detect significant effects at a respectable level of power. For Survey 1, I

consulted the raw frequencies Van Peer (1980: 499 ff.) reported in his underlining tests for strikingness. After confirming the *z* scores that Van Peer (1980: 292) reported, I calculated effect sizes for each of his stimulus texts using the formula d = z / sqrt(N), where d is the effect size, z is the *z* score, and N is the number of trials (the total number of encounters between readers and lines of poetry). In this way, I determined that the largest effect size in Van Peer (1980) was 0.57, while the smallest was 0.10. Thus, any result between these two extremes would not be unreasonable to expect in my own project, although setting the d value at 0.1 would result in a more sensitive design. Since I wanted to be able to compare responses from two groups (students and non-students), I next ran a power calculation for a two-sample t-test[33] in the statistical programme *R* to determine the number of trials I would need in each group. I set the d value at 0.1, with the alpha at 0.05 and the power at 0.8 per convention. According to this test, I would need to have enough respondents to result in 1570.73 (or 1571) trials *in each group* to be able to detect an effect size of 0.1.

However, since Survey 1.1 and Survey 1.2 each featured different translations, it was necessary to calculate a separate number of respondents for each. Dividing 1571 by the 56 lines presented in Survey 1.1[34] gave 28.05 (or 29 respondents needed in each group, i.e. 29 students plus 29 non-students). Dividing 1571 by the 52 lines of Survey 1.2 yielded 30.21 (or 31 respondents needed in each group). Thus 58 (29 x 2) respondents were needed for Survey 1.1, while 62 (31 x 2) were needed for Survey 1.2, for a total of 120. The same calculations applied to Survey 2, as it included a similar underlining task. Thus, a further 120 respondents were required. The total so far was 120 + 120 = 240.

---

[33] A non-parametric test would arguably have been more appropriate.
[34] As I explain in section 5.2.1, I counted tone-units instead of lines for Poem C.

Survey 3 would have added an additional 120 respondents; however, since I planned to compare Survey 3 responses to those collected from Survey 1, I would *already* have more than enough text-as-translation trials (or so I thought—see section 5.2.10). If 120 people had already filled out a text-as-translation survey in Survey 1 (58 people in Survey 1.1 and 62 in Survey 1.2), then this would already have resulted in 3142 trials (1571 x 2). According to my calculations in *R*, I would only need 1046.88 trials of the text-as-poem condition to detect an effect level of 0.1 at the conventional alpha and power. Dividing 1047 by the 56 lines of Survey 3.1 gave 18.70. Dividing 1047 by the 52 lines of Survey 3.2 gave 20.13. Thus, I would only need an additional 40 respondents (19 + 21). This would make the grand total 240 + 40 = 280 respondents.

I decided to recruit participants for Survey 1 first. To do this, I employed a systematic sampling technique (see Babbie 1990: 84-85) to sample from the phonebook probabilistically—probabilistic sampling being necessary in order to make statistical inferences beyond the samples themselves. However, it soon became clear that I would not meet the required number of respondents: the return rate for my surveys was dismal. Owing to the costliness of mailing surveys, I decided to supplement this original sampling method with a 'convenience' sampling method. Moreover, I decided to forego random sampling for Surveys 2 and 3, and to employ convenience sampling for these as well. I thus distributed Surveys 1, 2 and 3 to a variety of local groups to enlist non-student participants.

In addition to the classes mentioned above, I distributed Survey 1 to: several local poetry appreciation/writings groups, a local reading group, patrons of a local library, and a handful of acquaintances. I distributed Survey 2 to: a local reading group, a local adult language-learning group (not studying an Asian language), and several more acquaintances. I distributed Survey 3 to: another local poetry writing group, another

local library, and a weekly university gathering of students and teachers of a particular (non-Asian) language (although no responses from this group came back). Surveys were distributed between 16 May 2018 and 30 January 2019. Respondents were asked to return surveys using provided self-addressed, stamped envelopes.

Fig. 7a shows the number of respondents per group, per survey. However, as not all respondents responded to every text or survey item, these numbers may vary from hypothesis to hypothesis, and from text to text. Fig. 7b provides a further breakdown of respondents by sampling source.

Fig. 7a. No. Respondents per Survey

	Students (Undergraduates)	Non-Students	Total
Survey 1.1 (Texts A, B, C, D)	17	15	32
Survey 1.2 (Texts, E, F, G, H)	13	18	31
Survey 2.1 (Texts A, B, C, D)	7	5	12
Survey 2.2 (Texts E, F, G, H)	4	5	9
Survey 3.1 (Texts A, B, C, D)	0	4	4
Survey 3.2 (Texts E, F, G, H)	0	3	3
	41	50	91

Fig. 7b. No. Respondents per Sampling Group

	Level 1 Class	Level 2 Class	Level 3 Class	Mail (Random Sampling)	Convenience (Assorted)	Total
Survey 1.1	17	0	0	1	14*	32
Survey 1.2	13	0	0	2	16*	31
Survey 2.1	0	3	4	0	5	12
Survey 2.2	0	3	1	0	5	9
Survey 3.1	0	0	0	0	4	4
Survey 3.2	0	0	0	0	3	3
	30	6	5	3	47	91

*Includes 1 survey whose code was removed by the respondent

All in all, the response rate to my surveys was disappointingly low. In this sense, Surveys 2 and 3 might be better regarded as exploratory studies. ***The non-probabilistic nature of the sampling also leads to an important limitation: statistical inferences should not be made to populations beyond the respondents themselves.*** However, in this respect, my study is not unique—none of the previous studies cited in this project used properly random sampling, either. Moreover, unlike most of those studies, I have managed to enlist more non-student participants than student participants; in this respect, my samples might be thought to be more representative. Finally, as Miller (1984: 159) writes, researchers are "of course […] free to make discretionary judgements outside of the strict confines of scientific method, i.e. *non-statistical inferences*." Miller (ibid.: 160) explains that where non-statistical inferences are concerned, "the extent to which particular findings are held to be generally applicable is a matter for [the researcher] and the scientific community to decide, on the basis of what seems likely from our knowledge and intuitions about behaviour […]." Nevertheless, this remains a major limitation of the present study.

It will also be noted that no Japanese respondents were enlisted to respond to the STs. This was because my study was not designed to investigate the impact of translation *itself* on responses, but rather the potential impact of different translation styles, conceived here as either 'ostensibly domesticating' or 'ostensibly foreignising.' The foreknowledge that texts had been translated was indeed operationalised in Survey 3, but again, all respondents were TT readers. I ruled out recruiting ST respondents at an early stage chiefly out of time considerations: Van Peer (1986: 59) noted the "large amount of time involved in preparing full-scale stylistic analyses of even short texts," and I wanted to ensure that could I give my own analyses the time and attention they deserved.

To avoid bias, surveys were only opened and analysed *after all stylistics analyses were completed*. Next, survey responses were entered and organised in a password-protected Excel

file, and statistical analysis was performed using IBM's SPSS Statistics programme (version 25). The specific statistical techniques are discussed in section 5.2 under the relevant hypotheses.

4. STYLISTIC ANALYSES

4.1. SAMPLE ANALYSIS: BEICHMAN'S "SHIJIMI CLAMS"

4.1.1 PHONOLOGY

Below is my transcription of Beichman's "Shijimi Clams." This is the transcription I have used to analyse foregrounding on the phonological level of the poem. Like the rest of the transcriptions in this chapter, it accounts for weak forms and linking *r*'s.[35] Transcribing in a manner consistent with connected speech gives us the chance to identify potential FG that might otherwise escape detection in a phonological analysis of words in isolation.

ʃɪdʒɪmɪ klæmz

wəʊk ʌp ɪn ðə ded əv naɪt
ɪn ə kɔːnər əv ðə kɪtʃɪn
ðə lɪtl klæmz aɪd bɔːt ðæt iːvnɪŋ
wər əlaɪv maʊðz əʊpən

ət dɔːn
aɪl gɒbl jʊ ʌp
iːtʃ ən evrɪ wʌn

let aʊt ə kækl
laɪk ən iːvl əʊld wɪtʃ
ɑːftə ðæt kʊdnt help ɪt hæd tə
sliːp ɔːl naɪt wɪð maʊθ hɑːf əʊpən

Shijimi Clams

woke up in the dead of night—
in a corner of the kitchen
the little clams I'd bought that evening
were alive, mouths open—

"At dawn
I'll gobble you up
each and every one"

let out a cackle
like an evil old witch
after that couldn't help it had to
sleep all night with mouth half-open

---

[35] A few things to note about this transcription: first, I have used phonemic symbols from the A.C. Gimson (1962) system of notation. This is to facilitate comparison with Van Peer (1980). Second, the transcription is machine-assisted but manually checked: I first processed the text using the online transcription tool *toPhonetics* (https://tophonetics.com/), then to ensure accurate transcription, I cross-checked the results with the *OED Online* and other phonetics and phonology texts. In cases where more than one transcription was possible (for instance the word "open" /əʊp(ə)n/, where the schwa is optional), I have chosen the transcription that is closer to careful speech. This is because there is some cross-linguistic evidence to suggest that people tend to read poetry at a slower rate than other types of text (see, e.g., Fónagy and Magdics [1960], and Kowal and O'Connell [1980]).

The text is comprised of 58 orthographic words, including the title. The only punctuation marks used are two dashes and a set of quotation marks. Two other orthographic features: lines in this poem generally do not begin with capital letters, and two lines are completely italicised (lines 6 and 7). I discuss orthographic and graphological features under the heading of phonology because they so frequently suggest ways of reading words or phrases that distinguish them *phonologically* from the surrounding text (in much the same way that my italics here suggest that 'phonologically' should be read with extra emphasis). While it is not possible to say to a certainty what kind of intonational contour is denoted by the italics in lines 6 and 7, it seems likely from the context of the sentence that its intonation would have an 'attitudinal function.' This is the idea that intonation "is used to convey our feelings and attitudes: for example, the same sentence can be said in different ways, which might be labelled 'angry,' 'happy,' 'grateful,' 'bored,' and so on," as Roach (2009: 147) writes. While my analysis of intonation comes later, it shows falling intonation in both of the tonic syllables[36] in the italicised lines. As Roach (ibid.: 123) notes, the falling tone in English often gives the impression of finality. As the mood of the sentence is commissive, or at least resembles what in other languages is referred to as the commissive—that is, it expresses a "commitment to a future course of action" (Crystal 2008: 88)—it seems likely that the intonation of this sentence would express extra finality or emotion, comprising an internal deviation from the rest of the poem. Since there are two italicised lines, we may add this to our FG list as a parallelism. Insofar as these lines depart from the rest of the poem, which is un-italicised, we may also classify them as an internal deviation.

The same procedure applies to the use of punctuation in the poem. The two dashes comprise a parallelism. Since they break with the lack of punctuation of the rest of the poem

---

[36] The tonic syllable is the most prominent syllable in a tone unit. For this analysis, we may define a tone unit as an intonational structure consisting, maximally, of four different components: an optional pre-head (PH), an optional head (H), a requisite tonic syllable (TS), and an optional tail (T). For more on these terms, see Roach (2009: 130-133).

(apart from the two quotations marks), we can also classify them as an internal deviation. Likewise, the quotation marks in lines 5 and 7 break with the lack of punctuation in the rest of the poem (apart from the dashes). However, since quotation marks typically come in pairs, it does not make sense to describe their use here as a parallelism. What *is* parallel, however, is that the first and last lines of the first two stanzas all use punctuation of some kind.

A couple more things to note about orthography: the use of the capital letter at the beginning of lines 5 and 6 breaks the pattern established earlier (and resumed later) in the poem—I refer here to the pattern of *not* beginning lines with capital letters. Thus, it enters the inventory as an internal deviation. And insofar as *two* lines do this, those lines comprise their own parallelism.

Where the intonational structure of the text is concerned, we can add several more features to our FG list. The analysis below is prepared on the basis of Roach (2009) and Wells (2006), while its notation comes from Roach (2009). Tonic syllables are underlined, double vertical lines (‖) indicate a pause-type boundary, single vertical lines (|) indicate a non-pause boundary, while the ' mark represents a stressed syllable. The • symbol indicates a stressed syllable in the tail, and dashed lines (¦) represent boundaries between components within the tone units. Tones are represented as follows: falling (\), rising (/), fall-rise (∨), rise-fall (^), level (_). (Primary and secondary stress are not distinguished; neither are high and low heads. Finally, no stresses have been downgraded.)

```
    PH      TS          T
‖ Shi ┊ \ji ┊ mi • Clams ‖


    H      TS   PH      TS          T
'woke ┊ \up │ in the ┊ \dead ┊ of • night ‖
PH           H              TS          T
in a ┊ 'corner of the ┊ ∨kitch ┊ en ‖
PH        H      TS                T
the ┊ 'little ┊ ∨clams ┊ I'd • bought • that • evening │
PH          TS       H      TS      T
were   a ┊ \live ‖ 'mouths ┊ \o ┊ pen ‖


  PH      TS
At ┊ ∨ dawn ‖
PH         H          TS
I'll ┊ 'gobble you ┊ \up ‖
        H          TS
'each and 'every ┊ \one ‖


      H      TS      T
'let 'out a ┊ \cack ┊ le │
PH         H      TS
like an ┊ 'evil 'old ┊ \witch ‖
  H      TS      H      TS    T    H
'after ┊ ∨ that ‖ 'couldn't ┊\help ┊ it ‖ 'had to
          TS    PH      H      TS      T
'sleep 'all ┊ \night │ with ┊ 'mouth 'half ┊ \o ┊ pen ‖
```

197

The first thing to note is that all lines terminate in a boundary (either ‖ or |), except for line 10. This comprises an internal deviation. Next, there is a parallelism between the title and the first stanza—each line ends in the same pattern of a tonic syllable (TS) followed by a tail (T). A new pattern of line endings is then instantiated in the second stanza, whereby each line ends on a TS, comprising another parallelism.

One final thing to note about intonation and lineation: these two aspects of a poem can interact to influence the pronunciation of individual phonemes. An example of this occurs in line 10, already identified above as containing an internal deviation. Owing to the previous pattern of lines ending at tone-unit boundaries, this line initially appears to have the word "to" in final position, which would mean that it is pronounced something like /tu/ (Roach 2009: 93). However, since it comes before a consonant in a medial position in the sentence, its pronunciation in RP would reflect the weak form /tə/ (ibid.). This 'hiccup' would, in the theory of FG, be predicted to attract extra attention from the reader, so we may add it to our FG list. Van Peer (1980) classifies such ambiguities as statistical deviations, and I follow his precedent.

Related to intonation is the text's rhythmico-metrical make-up. If we understand metre as "a pattern of prominent and nonprominent, or marked and unmarked, positions that together form a distinct pattern" (Brogan 2012: 657), then we must confront the problem of what comprises a prominent or marked position. Reuven Tsur (2015: 172) diagnoses the problem succinctly when he remarks that "ictus" (the technical name for a marked position) actually refers to "an indefinite mixture of lexical stress and metrical downbeat." He goes on: "[…] students are confused—and you can't blame them for that—when they have to decide, in assigning prosodic marks, whether to mark a linguistic stress or metrical downbeat as ictus. When in the iambic metre, for instance, an unstressed syllable happens to be even-numbered in its sequence, they have no rule to decide whether to mark an ictus or a non-ictus" (ibid.).

For the sake of transparency, I have used lexical stress as my guideline for assigning ictus (indicated by the "–" mark).

Title	Stanza 1	Stanza 2	Stanza 3
x – x –	– – x x – x –	x –	– – x – x
	x x – x x x – x	x – x x –	x x – x – –
	x – x – x – – – x	– x – x –	– x – – x – x – x
	x x – – – x		– – – x – – – x

There is no overarching metrical pattern to the poem. That said, brief pieces of the poem are metrically regular: the title is iambic, line 2 repeats the figure x x – x, and line 11 comprises a pattern of three stressed syllables followed by one unstressed. The metrical figure of three stressed syllables in a row is known as a 'molossus.' As Cutler (2015: 107) notes, "[r]hythmically, English prefers to avoid successive stressed syllables." Thus we might also describe the lines containing molossi as statistically deviant. These are lines 3, 4 and 11. Moreover, both lines 3 and 4 end in an x – – – x pattern. Finally, the iambic pattern of the title is repeated at the beginning of line 3, while the trochaic pattern from the beginning of line 7 is repeated at the end of lines 8 and 10. These enter out FG list as examples of parallelism.

The next stage of the analysis concerns word length. In this analysis, word length is measured in letters. Grotjahn and Altmann (1993: 143) are quite right when they say that "the choice of the unit of measurement strongly affects the model of word length to be constructed." For instance, there can be "a large difference between the number of letters and the number of phonemes in a word" (ibid. 1993: 142). However, as Piantadosi et al. (2011: 3526) remark, orthographic length is "readily available from corpora and tends to be highly correlated with both phonetic length and production time." While Constable (1997) has some word-length data (measured in syllables) concerning English-language poetry, his corpus is smaller than the poetry subcorpus of the BNC and contains no text written after 1894.

Choosing orthographic word length has two benefits for this project: 1) it is much less time-consuming to calculate orthographic as opposed to syllabic word-length frequencies in the BNC, and 2) orthographic length is immediately visible on the page, so if there is a visual component to word length as it relates to FG, then this will be partly reflected in the analysis.

I used AntConc 3.5.7 (Anthony 2018) to determine word-length frequencies for the poetry subcorpus of the BNC. This is reflected in the table below. Note: 219,904 words were counted. Apostrophes and hyphens within words do not contribute to the letter-count. Thus, "couldn't" is a seven-letter word, and "half-open" is an eight-letter word.

Fig. 8. BNC Word-Length Frequencies: Poetry Subcorpus

N-letter word	Freq	Percentag
1	9602	4.37%
2	34730	15.79%
3	47625	21.66%
4	40792	18.55%
5	28541	12.98%
6	19484	8.86%
7	15873	7.22%
8	9742	4.43%
9	6300	2.86%
10	3652	1.66%
11	1864	0.85%
12	879	0.40%
13	470	0.21%
14	191	0.09%
15	90	0.04%
16	35	0.02%
17	15	0.01%
18	8	0.00%
19	4	0.00%
20	1	0.00%
21	2	0.00%
22	1	0.00%
23	1	0.00%
24	2	0.00%
Total no. of words=	219904	100.00%

From the above we see that three-letter words are used most frequently in the poetry subcorpus of the BNC. In general, the longer a word is, the less likely it is to appear in that corpus.

The question now is: how does Beichman's "Shijimi Clams" stack up in terms of word length? According to the table below, Beichman's translation seems to overperform in some categories (see, e.g., 5-letter words) and underperform in others (see, e.g., 8-letter words).

Fig. 9. Poem A: Distribution of Words by Length

	1-Letter	2-Letter	3-Letter	4-Letter	5-Letter	6-Letter	7-Letter	8-Letter	Total words per line
Title					1		1		2
1		3	1	2	1				7
2	1	2	1			1	1		6
3		1	1	1	1	2	1		7
4				2	1	1			4
5		1		1					2
6		1	2			1			4
7			2	1	1				4
8	1		2			1			4
9		1	1	2	1				5
10		2	1	2	1		1		7
11			1	1	3			1	6
12									
13									
14									
15									
16									
Total	2	11	12	12	10	6	4	1	58
Observed %	3.45%	18.97%	20.69%	20.69%	17.24%	10.34%	6.90%	1.72%	
Expected %	4.37%	15.79%	21.66%	18.55%	12.98%	8.86%	7.22%	4.43%	

However, to determine whether these differences in percentage between observed and expected values are statistically significant, we will need to perform either a chi-square goodness-of-fit test or a log-likelihood goodness-of-fit test. I have chosen the log-likelihood method (also known as a *G*-test) because there are expected frequencies of less than five in the contingency table used for calculating the chi-square goodness-of-fit. In cases like this, the *G*-test "provides a better testing procedure" (Mellinger and Hanson 2017: 175).

Results from this calculation, using McDonald's (2014) *G*-test of goodness-of-fit spreadsheet, reveal that in fact the observed distribution is consistent with the expected one,

such that G = 9.887, p (two-tailed) = 0.872, and df = 16.[37] In other words, the word-length frequencies in Beichman's translation are not unusual.

Nevertheless, words of different lengths are not evenly distributed throughout the text. Line 11 has three times the amount of 5-letter words as any other line. Meanwhile, the text's 3-letter words are concentrated in lines 6, 7 and 8. Finally, there is a concentration of 2-letter words in lines 1 and 2. We may add these features to our FG list as statistical deviations.

Next is an analysis of individual phonemes. The ideal reference corpus would be the same corpus used above. However, as a phonemic analysis of that amount of text would not be feasible, I have opted to use the phoneme frequencies given in Knowles (1987: 223-224). While Knowles' phoneme frequencies are very similar to those given by Fry (1947),[38] there are three advantages to using Knowles' frequencies: 1) they are more recent, and therefore probably closer to actual contemporary values; 2) they are based on ten different types of text, while Fry's (1947) data came from only one text; and 3) Knowles provides the total number of phonemes counted, which makes a log-likelihood calculation possible. As we saw above, this number tells us whether the difference between observed values and expected values is statistically significant. In this instance, since I am analysing individual phonemes, I am using a test of independence rather than a test of goodness-of-fit.

These log-likelihood values are displayed in the right-most column (LL) of the tables below. The first table tallies the vowels used in the poem. The second table tallies consonants. The third table organises the sounds in the poem according to articulatory criteria. The classification of vowels and consonants in the third table is based on Van Peer (1980); no claim is put forward as to the appropriateness of these classifications— they are adopted here to make comparison with Van Peer (1980) more convenient. In each of the three

---

[37] For this calculation, I have used the first 17 categories from the word-length frequency table discussed earlier. This is because the percentages of categories 18 to 24 are so miniscule. At two significant figures, the percentages for the first 17 categories already add up to 100%.

[38] These are the frequencies used in Van Peer (1986).

tables, Fry's (1947) values are included for the sake of thoroughness, but the column for

observed-over-expected values (Obs/Exp) uses Knowles' values.

Fig. 10. Vowel/Diphthong Frequencies

Line	Title	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total	Obs %	Fry	Knowles	Obs/Exp	LL
ə		2	4	1	3	1		1	1	1	2	1						17	9.50%	10.74	10.49	0.91	-0.17
ɪ	3	1	3	2			1		1	1	1	1						13	7.26%	8.33	8.26	0.88	-0.22
e		1					1	1		1								4	2.23%	2.97	2.57	0.87	-0.08
ʌ		1				1	1											3	1.68%	1.75	1.41	1.19	0.08
i:			1				1		1		1							4	2.23%	1.65	1.80	1.24	0.17
æ	1		2				1		2									6	3.35%	1.45	1.80	1.86	1.86
ɒ					1													1	0.56%	1.37	1.73	0.32	-1.91
ɔː			1	1		1					1							4	2.23%	1.24	1.36	1.64	0.82
uː																		0	0.00%	1.13	1.46	0.00	-5.18
ʊ						1			1									2	1.12%	0.86	0.38	2.94	1.63
ɑː									1	1								2	1.12%	0.79	0.56	2.00	0.75
ɜː																		0	0.00%	0.52	0.62	0.00	-2.2
aɪ		1		1	1		1		1		1							6	3.35%	1.83	2.22	1.51	0.87
eɪ																		0	0.00%	1.71	1.54	0.00	-5.46
əʊ		1		1				1			1							4	2.23%	1.51	1.59	1.41	0.41
aʊ				1			1				1							3	1.68%	0.61	0.65	2.58	1.96
ɛə																		0	0.00%	0.34	0.31	0.00	-1.1
ɪə																		0	0.00%	0.21	0.36	0.00	-1.28
ɔɪ																		0	0.00%	0.14	0.26	0.00	-0.92
ʊə																		0	0.00%	0.06	0.04	0.00	-0.14
Total V per line	4	7	8	8	6	2	4	5	4	5	8	8	0	0	0	0	0	69	38.55%	39.21	39.41		
Total phonemes in poem	179																						

Fig. 11. Consonant Frequencies

Line	Title	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total	Obs %	Fry	Knowles	Obs/Exp	LL
n		2	3	1	1	1		2		1	1	2						14	7.82%	7.58	7.65	1.02	0.01
t		1		3		1		2			5	1						13	7.26%	6.42	7.48	0.97	-0.01
d		2		1		1			1	2								7	3.91%	5.14	4.12	0.95	-0.02
s												1						1	0.56%	4.81	4.77	0.12	**-10.67**
l	1			3	1		2		2	3		3						15	8.38%	3.66	3.91	2.14	**6.71**
ð		1	1	2	1						1	1						7	3.91%	3.56	3.37	1.16	0.14
r			1			1			1									3	1.68%	3.51	3.62	0.46	-2.31
m	2				1	1						1						5	2.79%	3.22	2.29	1.22	0.18
k	1		1	2	1				2	1	1							9	5.03%	3.09	2.89	1.74	2.26
w			1			1		1		1		1						5	2.79%	2.81	2.53	1.10	0.05
z	1				1	1												3	1.68%	2.46	3.05	0.55	-1.31
v		1	1		1	1			1		1							6	3.35%	2.00	1.94	1.73	1.48
b				1		1												2	1.12%	1.97	2.17	0.51	-1.1
f										1		1						2	1.12%	1.79	0.66	1.69	0.46
p		1			1		1				1	2						6	3.35%	1.78	2.05	1.64	1.21
h										2	1							3	1.68%	1.46	1.00	1.68	0.66
ŋ				1														1	0.56%	1.15	0.94	0.59	-0.32
g							1											1	0.56%	1.05	0.93	0.60	-0.31
ʃ	1																	1	0.56%	0.96	0.82	0.68	-0.17
j							1											1	0.56%	0.88	1.26	0.44	-0.87
dʒ	1																	1	0.56%	0.60	0.63	0.89	-0.01
tʃ			1					1		1								3	1.68%	0.41	0.53	3.16	2.73
θ												1						1	0.56%	0.37	0.57	0.98	0
ʒ																		0	0.00%	0.10	0.04	0.00	-0.14
Total C per line		7	10	9	16	9	3	6	6	6	9	14	15	0	0	0	0	0	110	61.45%	60.78	59.22	1.04	
Total phonemes in poem	179																							

Fig. 12. Total Sound Inventory

Total Sound Inventory			No.	Obs %	Fry	Knowles	Obs/Exp
Vowels			69	31.28	32.80	32.44	0.96
Diphthongs			13	7.26	6.41	6.97	1.04
Consonants			110	61.45	60.78	59.22	1.04
Vowels							
			No.	Obs %	Fry	Knowles	Obs/Exp
Close	(iː, ɪ, uː)		17	9.50	11.11	11.52	0.82
Half-Close	(ə, ɜː, e, ʊ)		23	12.85	15.09	14.06	0.91
Half-Open	(æ, ʌ, ɒ, ɔː)		14	7.82	5.81	6.3	1.24
Open	(ɑː)		2	1.12	0.79	0.56	2.00
Frontal	(iː, e, æ, ɪ)		27	15.08	14.4	14.43	1.05
Central	(ʌ, ɜː, ə, ʊ, ɑː)		24	13.41	14.66	13.46	1.00
Back	(ɒ, ɔː, uː)		5	2.79	3.74	4.55	0.61
Consonants							
			No.	Obs %	Fry	Knowles	Obs/Exp
Semi-Vowels	(w, j)		6	3.35	3.69	3.79	0.88
Liquids	(l, r)		18	10.06	7.17	7.53	1.34
Nasals	(m, n, ŋ)		20	11.17	11.95	10.88	1.03
Fric. & Asp.	(f, θ, s, ʃ, v, ð, z, ʒ, h)		24	13.41	17.51	16.22	0.83
Affric.	(tʃ, dʒ)		4	2.23	1.01	1.16	1.93
Plosives	(p, t, k, b, d, g)		38	21.23	19.45	19.64	1.08
Labials	(m, p, b, w)		18	10.06	9.78	9.04	1.11
Dentals	(f, v, θ, ð)		16	8.94	7.72	6.54	1.37
Alveolars	(t, d, s, z, n, l, r)		56	31.28	33.58	34.6	0.90
Palatals	(tʃ, dʒ, ʃ, ʒ, j)		6	3.35	2.95	3.28	1.02
Velars	(k, g, ŋ)		11	6.15	5.29	4.76	1.29
Glottals	(h)		3	1.68	1.46	1.00	1.68

204

As we can see from the tables above, there are only four cases where the difference between the observed and the expected values is statistically significant. These cases are highlighted in yellow in the LL column, as their absolute values are past the 3.8 threshold for statistical significance at $p < 0.05$. First we have the two vowels /uː/ and /eɪ/. Their negative LL value indicates underuse in the text—they are so underused, in fact, that they do not appear at all in Beichman's translation. Since they do not appear, there is no opportunity for them to be underlined by readers. Thus, we must forgo adding them to our FG list. Where consonants are concerned, however, it is possible to add two items: /s/ and /l/. While /s/ is significantly underused, it *does* appear in the poem, so we may add it to our list as a statistical deviation. Meanwhile, /l/ is overused, so we may add this as well.

Just because the other speech sounds do not have a high LL value does not mean they may not be predicted to contribute to FG. Some of these sounds are more highly concentrated in certain parts of the text. Line 4 has four of the poem's /ə/ sounds. The vowel /ɪ/ is concentrated in the title and line 2. There are five instances of /t/ in line 10. The /l/ is concentrated in lines 3, 9 and 11. Finally, there are three /n/ sounds in line 2. We may add all these as statistical deviations.

There is nothing in the way of obvious end rhyme, but "kitchen" and "witch" comprise a near rhyme. There are also instances of words sharing vowel sounds in proximity to one another. Thus, there is the /ae/ assonance between "clams" and "that" in line 3, as well as between "cackle," "that" and "had" in the third stanza. There is the /aɪ/ assonance between "night," "I'd" and "alive" in the first stanza. There is the /ɔː/ assonance between "corner" and "bought" in lines 2 and 3. There is line-final /ʌ/ assonance between "up" and one" in lines 6 and 7. The rest of the above instances enter our list as parallelisms.

There are also several words sharing consonants in close proximity. For instance, there is the /k/ alliteration between "corner," "kitchen" and "clams" in lines 2 and 3. Then there is the

/l/ alliteration between "let" and "like" in lines 8 and 9. Then, near the end of the poem, we encounter the /h/ alliteration between "help," "had" and "half-open" (lines 10 and 11).

Finally, we might look at the distribution of consonant clusters in the poem. Unfortunately, analysing the frequency of consonant clusters in the poetry subcorpus of the BNC is beyond the scope of this study. Thus, we will not be able to compare the current text to a reference corpus. However, we can still make observations about how evenly (or unevenly) such words are distributed throughout the poem. The table below lists all consonant clusters by line number. As there is often "considerable" disagreement between speakers about syllable-counts (Roach 2009: 56), my inventory of consonant clusters does not take syllable divisions into account.

Fig. 13. Consonant Cluster Frequencies

	C_2-clusters	C_3-clusters	C_4-clusters	Total		
Title	2			2	/kl/ /mz/	
1						
2						
3	4			4	/tl/ /kl/ /mz/ /vn/	
4	1			1	/ðz/	
5						
6	1			1	/bl/	
7	1			1	/vr/	
8	1			1	/kl/	
9	2			2	/vl/ /ld/	
10	2	1		3	/ft/ /lp/	/dnt/
11						
12						
13						
14						
15						
16						
Total	14	1		15		

As we can see from the above, C_2 clusters are particularly concentrated in line 3. We may add this to our FG list as a statistical deviation. More consonant clusters end in /l/ than in any other sound. Since there is one in every unit of the poem set off by white space (i.e., the title and each stanza), we may classify this as a kind of pattern. Meanwhile, line 10 contains the text's only C_3 cluster. We may add this last instance as an internal deviation.

Perhaps the final thing to say about the phonology of the poem concerns the word "Shijimi." The question of how to analyse non-native words is an interesting one, both in terms of level of analysis (i.e., phonology, syntax or semantics) as well as type of deviation (internal, statistical or determinate). It is not a problem encountered by Van Peer (1980) in his stylistic analyses. Nevertheless, we would probably predict the item *shijimi* to attract at least *some* extra attention from participants, since it is unlikely that this word featured in the mental lexicons of readers prior to encountering the text. Results from eye-tracking studies (see Williams and Morris 2004; and Wochna and Juhasz 2013) show that readers engage in longer initial processing of novel words. It is possible that a reader's gaze duration "may index an attempt to integrate the meaning of the newly fixated word with the sentence context" (Wochna and Juhasz: 2013: 358), suggesting an "attempt, and failure, to access" the novel word from the reader's mental lexicon (ibid.). However, rather than include *shijimi* in the semantics section of my analysis, I have chosen to analyse it in the phonology section for two reasons. The first concerns the "incremental" way adults acquire new words (see Leach and Samuel 2007: 307): since the process begins here at the orthographic input stage, it seemed appropriate to discuss the phenomenon under the heading of phonology. The second reason concerns the phonotactics and spelling of 'shijimi.' The consonant-vowel structure of this word is CVCVCV. It is not difficult to think of a few words of this type in English off the top of one's head: happily, potato, tomorrow, gorilla, domino, etc. However, CVCVCV words featuring *the same vowel* in every syllable are harder to think of. 'Banana' would count in terms of spelling, but its pronunciation in RP involves two different vowels: /bənɑːnə/. Moreover, CVCVCV words ending with a vowel in the vicinity of /ɪ/ or /iː/ are not usually spelled with a word-final "i"—apart from borrowings from other languages (e.g. *chupatti*). The spelling of *shijimi* thus arguably marks it out as a loan as opposed to a phonologically possible, but heretofore non-existent English word. Because its phonotactics are possible in

English, it would be difficult to describe the word in terms of determinate deviation (a word beginning with /ŋ/, say, would violate English phonotactic constraints). So: should we classify it as an instance of internal deviation, insofar as it is the only loan in the translation? Or should we call it a statistical deviation, insofar as its particular CVCVCV structure (same vowel throughout) and spelling pattern seem relatively rare in English? I see no reason why it could not attract attention for both reasons, and have entered it into our FG list accordingly.

Statistical Deviation

- ambiguity concerning pronunciation of "to" in line 10

- molossus in lines 3, 4, 11 (x 2 in line 11)

- concentration of 5-letter words in line 11

- concentration of 3-letter words in lines 6, 7, 8

- concentration of 2-letter words in lines 1, 2

- /s/ underused by comparison with reference corpus (line 11)

- /l/ overused by comparison with reference corpus (title, lines 3-4, 6, 8-9, 11)

- concentration of /ə/ in line 4

- concentration of /ɪ/ in title and line 2

- concentration of /t/ in line 10

- concentration of /l/ in lines 3, 9, 11

- concentration of /n/ in line 2

- concentration of C_2 clusters in line 3

- particular CVCVCV structure of "Shijimi" (title)

Determinate Deviation

- none

Internal Deviation

- one loanword ("Shijimi") (title)

- attitudinal function of intonation in stanza 2 (lines 6-7)

- use of italics in lines 6, 7

- use of dashes in lines 1, 4

- uses of quotation marks in lines 5, 7

- lines 5 and 6 begin with capital letter

- all lines end at tone-group boundary, except 10

- iambic pattern in title

- repeated x x – x pattern in line 2

- regular metre in line 11

- line 10 contains only C_3 cluster

Parallelism

- two italicised lines (6 and 7)

- two dashed lines (1 and 4)

- first and last lines of stanzas 1 and 2 use punctuation (lines 1 & 4, 5 & 7)

- two lines (5 and 7) begin with capital letter

- each line in second stanza ends on a tonic syllable (lines 5-7)

- title and each line in first stanza end in TS, T pattern (title, lines 1-4)

- each line in second stanza ends in TS (lines 5 -7)

- iambic pattern in title and start of line 3

- trochaic pattern in parts of lines 7, 8, 10

- lines 3 and 4 end in an x – – – x pattern

- /ae/ assonance in line 3 ("clams," "that")

- /ae/ assonance in Stanza 3 ("cackle," "that," "had,") (lines 8, 10) (x 2 in line 10)

- /aɪ/ assonance in Stanza 1 ("night," "I'd" and "alive") (lines 1, 3, 4)

- /ɔ:/ assonance in lines 2 and 3 ("corner," "bought")

- line-final /ʌ/ assonance in lines 6 and 7 ("up," "one")

- near-rhyme between "kitchen" and "witch" (lines 2, 9)

- /k/ alliteration in lines 2, 3 ("corner," "kitchen," "clams") (lines 2, 3) (x 2 in line 2)

- /l/ alliteration in lines 8, 9 ("let," "like")

- /h/ alliteration in lines 10, 11 ("help," "had," "half-open")

- consonant clusters ending in /l/ in title + lines 3, 6, 8, 9 ("clams," "little," "clams," "gobble," "cackle," "evil")

Beichman's "Shijimi Clams" consists of six sentences, which I have diagrammed below with the aid of the Stanford Parser (Klein and Manning 2003). I have generally used the same phrase structure rules and labels used in Carnie (2013), with a few exceptions based on the Penn Treebank II Tag Set (see: CLiPS 2018). This is an effort to keep the trees from becoming too convoluted. To label the resulting syntax trees, I used Shang's (2011) Syntax Tree Generator. The Ø symbol in Sentences 1, 2, 4, 5 and 6 represents a deletion at the surface level of the sentence. The label S in S_6 indicates an infinitival clause.

Fig. 14. Sentence 1 (S_1)

```
                    TP₁
          ┌──────────┴──────────┐
         NP                     VP
          │              ┌───────┴───────┐
          N             V'               PP
          │           ┌──┴──┐         ┌───┴────┐
          Ø          V      P         P        NP
                     │      │         │     ┌───┼────┐
                   woke    up        in     D   N    PP
                                           │    │   ┌─┴─┐
                                          the  dead P   NP
                                                    │    │
                                                   of    N
                                                         │
                                                       night
```

Fig. 15. Sentence 2 (S_2)

```
                                      TP₂
        ┌──────────────────┬───────────────────────────────┐
       PP                  NP                               VP
    ┌───┴──┐       ┌────────┼──────────┐           ┌────────┼─────────┐
    P      NP      D       AdjP         N'          V       AdjP       TP₄
    │   ┌──┴──┐    │        │       ┌────┴───┐      │        │      ┌───┴───┐
   in   D     N'  the      Adj      N       TP₃   were      Adj    NP      VP
        │   ┌─┴─┐           │       │     ┌──┴──┐            │    ┌─┴─┐   ┌─┴──┐
        a   N   PP        little  clams  NP  T   VP        alive  D   N   V  AdjP
            │  ┌─┴─┐                      │   │ ┌─┴─┐             │   │   │   │
         corner P  NP                     N  'd V   NP           Ø mouths Ø  Adj
               │  ┌─┴─┐                    │    │   │                        │
              of  D   N                    I  bought D  N                   open
                  │   │                           │  │
                the kitchen                    that evening
```
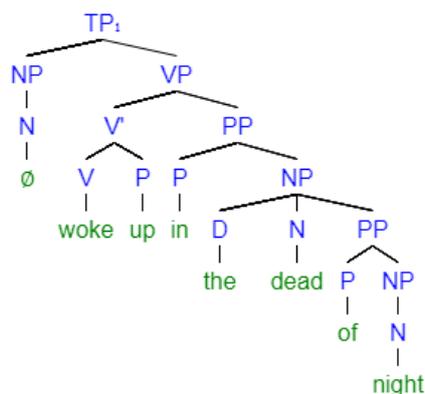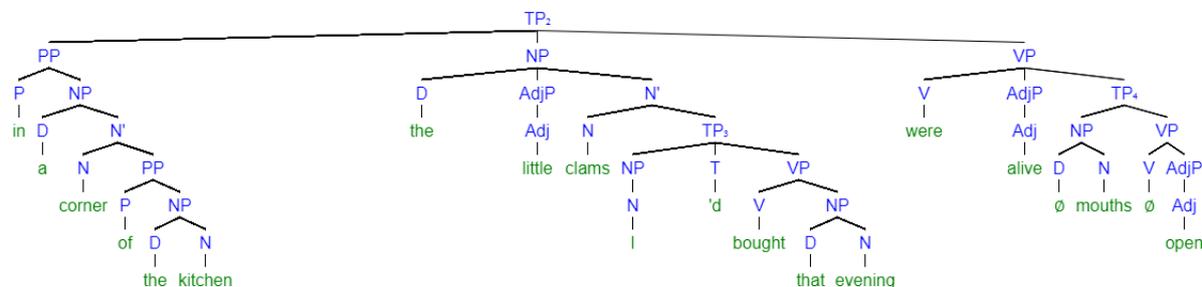
Fig. 16. Sentence 3 (S₃)
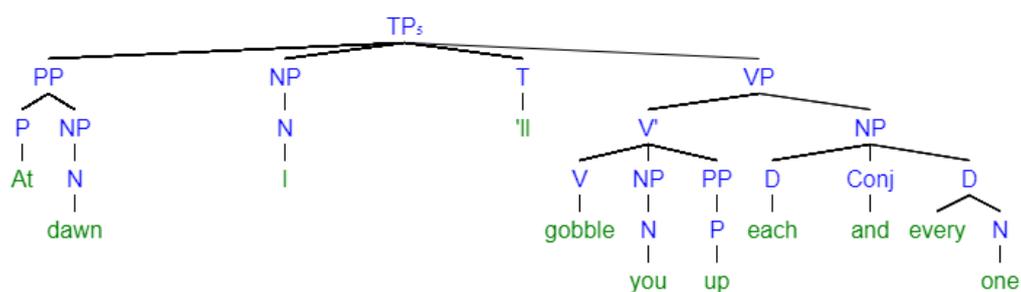
Fig. 17. Sentence 4 (S₄)
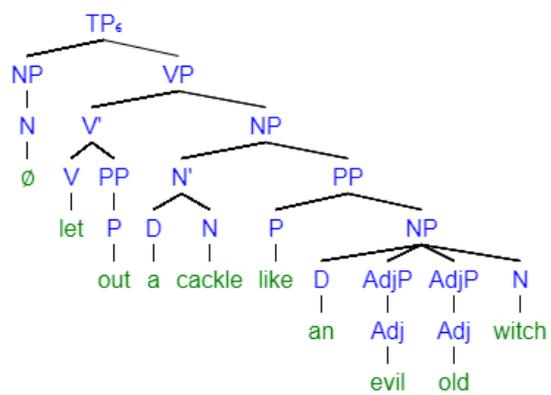
Fig. 18. Sentence 5 (S₅)
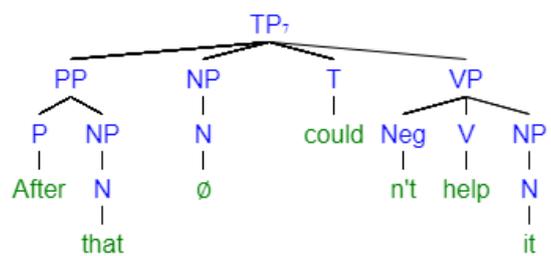
Fig. 19. Sentence 6 (S$_6$)
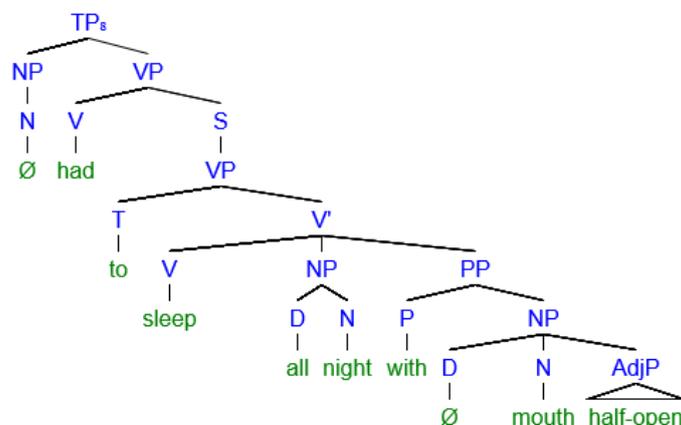
In the case of Sentences 1, 4, 5 and 6 (these correspond to lines 1, 8 and 10) there is a type of ellipsis called "Main Subject Deletion" (Gandón-Chapela 2013: 295). While at first glance this may seem an instance of determinate deviation, I have entered it in the FG inventory under the heading of statistical deviation. This is because the deleted element ("I") does not contravene the syntactic rules for ellipsis identified by Quirk et al. (1985: 896-897). In other words, it still produces a sentence acceptable to native English speakers.[39] Moreover, since the deletion of the same main subject occurs four times in the poem, this may be said to comprise a pattern. Thus, this feature also belongs in our FG inventory as an instance of parallelism. At the same time, the explicit appearance of "I" in S$_2$ and S$_3$ (lines 3 and 6) would seem to break this pattern of expectation. While it is likely that the deletion of "I" from these two sentences *in their current form* would make them less syntactically well-formed, these are not the only constructions the translator could have used. In other words, it would not have been impossible to continue the pattern of Main Subject Deletion. Thus, I have entered the two "I"s as an instance of internal deviation. Meanwhile, the occurrence of these two "I"s also comprises a kind of parallelism.

---

[39] As Teddiman (2011: 72) writes, cases of situational ellipsis "should not be classed as errors because they are both fully interpretable and not uncommon in casual speech."

The Main Subject Deletion in line 1 also creates a potential structural ambiguity. This is the possibility of initially reading the title as the subject of the first sentence. This reading is enabled by the fact that the first word in line 1 is not capitalised. Van Peer (1980) classifies structural ambiguity as a statistical deviation; I have followed his example.

I have mentioned that S_2 contains an instance of Main Subject Deletion. But it also contains two other kinds of ellipsis. First, it contains a verbless clause (see TP_4 in the diagram; the corresponding line in the poem is 4). As Quirk et al. (1985: 996) write, it is "usually possible to interpret [such clauses] as having an omitted BE […]." Thus, we can think of the construction "mouths open" at a more abstract level as "mouths were open." At the same time, we can also think of "mouths open" as "their mouths open," with the possessive restored. Thus, we can enter two more instances of statistical deviation: ellipsis of verb and ellipsis of possessive in line 4. S_6 (line 11) shares the same kind of possessive ellipsis: "with [my] mouth half-open" (however, since it does not comprise its own clause, but is instead simply a prepositional phrase, we cannot say that a BE verb has been omitted). The two possessive ellipses comprise another parallelism for our FG inventory. Additionally, S_2 (line 4) and S_6 (line 11) both end with the same pattern of a noun plus postpositive adjective.

Another structural similarity can be found in the prepositional phrases in S_1 and S_2 (see lines 1 and 2). Both phrases contain the same P D N P NP pattern, as the table below shows:

	P	D	N	P	NP
1.	in	the	dead	of	night
2.	in	a	corner	of	the kitchen

This enters the FG inventory under the heading of parallelism.

Two further parallelisms concern the position of time adverbials in the poem. In the first stanza, the two time adverbials ("in the dead of night" and "that evening") occur *after* their respective verbs, comprising one parallelism. However, this pattern is broken by the next two time adverbials, which are both fronted ("At dawn" and "after that"), comprising their own parallelism as well as a kind of internal deviation.

We can identify a number of other FG features. First is the high incidence of prepositional adverbs. These are words that look like prepositions (e.g. up, off, and out) but which modify verbs. Leech (1995: n.p.) provides two examples: "Come *out* here" and "I can't hold *out* any longer." (While *out* is considered a prepositional adverb in each example, it is also part of a phrasal verb in the second instance, since it changes the meaning of the verb.) Beichman's translation contains "woke up," "gobble […] up" and "let out" (see lines 1, 6 and 8). As there is one prepositional adverb per stanza, we may say this comprises a pattern, and so enter it as a parallelism. A search for prepositional adverbs in the poetry subcorpus of the BYU-BNC reveals that these occur at a rate of 10,665.01 per million, or 1.07 percent. Meanwhile, in Beichman's translation, prepositional adverbs occur at a rate of 3 in 58 words, or 5.17 percent. To see whether the difference between these frequencies is statistically significant, we need a statistical test that involves a *p*-value. When we use Rayson's (2016) Log-Likelihood and Effect Size Calculator, we see that the prepositional adverbs in Beichman's translation have a keyness value of 4.71. This result is significant at $p < 0.05$. This means there are more prepositional adverbs in Beichman's text than would be expected, considering the text type. This enters our list as a statistical deviation.

Next, we will look at the tense, aspect and mood of the verbs. To begin with, all the sentences in the poem are in the past tense, except for one. The exception is the future tense in S_3 (lines 5 to 7), which comprises an internal deviation. Similarly, all verbs in the poem are unmarked for aspect, except for one: the 'had bought' in S_2's (lines 2 to 4) "I'd bought,"

which is in the perfect aspect and comprises another internal deviation. Finally, all sentences in this poem are in the indicative mood except for two. S_3 (lines 5 to 7) resembles what in other languages is referred to as the commissive mood, whereby a speaker "makes a commitment to a future course of action" (Crystal 2008: 88). Meanwhile S_6 (lines 10 to 11) is in the obligative mood, which "signals the speaker's estimation of the necessity that the proposition expressed in his or her utterance be brought about" ("Obligative Mood" 2003: n.p.). Both comprise internal deviations.

Where other structures are concerned, there are several more features worth noting. One is a potential structural ambiguity involving lines 9 and 10. Due to the lack of end punctuation, it is possible to initially read S_4 as spilling over from line 9 into line 10, yielding the sentence "let out a cackle like an evil old witch after that." As I mentioned earlier, this kind of ambiguity is classified by Van Peer (1980) as a statistical deviation. Meanwhile, the lineation of this stanza creates an internal deviation as well. Whereas all other lines in this poem terminate at a more or less major constituent boundary, line 10 breaks that pattern. That is, one might expect line 10 to end after "it," which is where a period would normally go. Instead, we have the start of another sentence, and the line ends in enjambment. For the first time in the poem, the syntactic and typographical patterns do not coincide. This may attract extra attention from the reader.

Next, there is a deviation concerning the phrase "each and every one" in S_3 (line 7). While the sentence is completely grammatical, the position of the quantifiers within it recalls a phenomenon known as "quantifier floating." This term refers to the rightward movement of a quantifier from its typical position in a sentence, which would be nearer to the phrase it modifies. In this analysis, a more typical position for "each and every" would be in the main clause, nearer to the phrase it modifies, i.e.: "At dawn, I'll gobble each and every one of you up." However, as a rightward shift within the main clause would affect the acceptability of

the sentence (e.g. ?I'll gobble you each and every one up), the phrase "each and every" is set off syntactically, forming an adjunct. Insofar as rightward movement (like the examples of ellipsis discussed above) represents a change at the surface level of a sentence rather than a violation of any syntactic rule, it should enter our FG inventory under category of statistical deviation.

In terms of syntactic complexity, the poem arguably reaches its peak in S_2 (lines 2-4), which contains the only relative clause in the poem ("[that] I'd bought that evening"), in addition to an adjunct ("mouths open").[40] In a poem with so many simpler syntactic structures, this may be seen as an internal deviation. Moreover, the position of the relative clause between the subject of the sentence ("clams") and the main verb ("were") comprises what is known as a grammatical hiatus, whereby a break or gap is created between two closely related words by the insertion of "any number of linguistic phenomena (interjections, appositions, subordinate clauses, parenthetical remarks, etc.)" (San 2005: 137). As San (2005: 138) writes, subject-verb hiatus can be a particular source of "strain" for readers. While the subject-verb hiatus here is a short one, it nonetheless delays the onset of the verb and makes the syntax more complex. Since it is the only such hiatus in the poem, it is perhaps best described as an internal deviation.

Insofar as there are no coordinating or subordinating conjunctions in the poem, we would be justified in calling its syntax 'paratactic'—at least according to some definitions of the term (see e.g. Hanssen 2002: 140, Crystal 2008: 350). In the above-cited works, parataxis is characterised as the absence of *any and all* conjunctions (whether coordinating or subordinating). But sometimes the use of coordinating conjunctions is described as paratactic

---

[40] Empirical research supports this characterisation. Isakson and Spyridakis (2003: 539-40) sum up the findings of previous research: "Readers recall independent clauses […] faster than dependent clauses […] (Townsend, Ottaviano, and Bever 1979). […] Larkin and Burns (1977) similarly found decreased recall and comprehension of information in embedded relative clauses." S_2 is the only sentence in this text with an embedded relative clause.

(see Quartermain [2013: 47]). In the latter understanding of the term, the important thing is that the structures linked are co-ordinate, that is, of equal syntactic status. For the purposes of this analysis, I will take 'paratactic' to mean the juxtaposition of co-ordinate sentences or clauses with or without coordinating conjunctions.

The opposite of parataxis is hypotaxis. This refers to the use of "dependent constructions, especially those where constituents have been linked through the use of subordinating conjunctions" (Crystal 2008: 233). By this definition, the relative clause in S_2 (line 3) ("[that] I'd bought that evening") is indeed hypotactic, since it is a kind of subordinate clause. The use of this subordinate clause would therefore comprise a deviation from the overriding pattern of parataxis. We may enter this as an instance of internal deviation. Meanwhile, the background pattern of parataxis comprises a kind of parallelism,

Finally, I used an online corpus analysis tool known as Wmatrix (Rayson 2009) to try to determine whether any parts of speech (POS) might be predicted to stand out on the basis of their keyness value. This procedure is to be distinguished from the keyness analysis undertaken in the semantics section. In that section, the keywords analysis helps us to predict the strikingness of individual lexemes with no consideration for their part of speech. By contrast, the key POS analysis here only singles out individual lexemes insofar as they belong to *syntactic categories* that are over- or underrepresented in the target text when compared to the reference corpus. The assumption is that readers approach the text with subconscious expectations concerning not just the relative frequencies of different kinds of semantic content (as manifested in lexis), but also with expectations surrounding the distribution of different POS; moreover, these expectations are formed by previous experience with the text type in question. Since it is not feasible to tailor strikingness predictions according to each individual reader's experience with poetry, a reference corpus designed to be representative of that text-type is employed. The usefulness of key POS analysis lies in identifying items

that might otherwise be overlooked in a keywords analysis but which may nonetheless attract extra attention by belonging to an under- or overrepresented POS category. For my reference corpus, I uploaded the same BNC poetry subcorpus which I use in my semantics analyses. While no POS were identified as key at a significance level of $p < 0.01$ (corresponding to a keyness value of 6.6 or higher), three categories were flagged as key at $p < 0.05$ (having a keyness value of at least 3.8). A significance level of $p < 0.05$ is normally considered acceptable.[41] These categories were: VHD (past tense forms of HAVE), DD1 (singular determiners), and DB (determiners capable of functioning as pronouns). According to the Wmatrix analysis, VHD refers to *'d* and *had* (lines 3 and 10), DD1 to *that* (lines 3 and 10) and *each* (line 7), and DB to *all* (line 11) and *half* (line 11). We can describe the first two as statistical deviations. However, it appears that *half* has been incorrectly identified; it is used in the poem as an adverb, not a determiner. Since the DB category would otherwise not have been identified as key, I will not add this category to the FG list.

Statistical Deviation

- main subject deletion in lines 1, 8, 10 (x 2 in line 10)

- structural ambiguity involving title and subject of line 1

- ellipsis of be-verb in line 4

- possessive ellipsis in lines 4, 11

- high occurrence of phrasal verbs (see lines 1, 6, 8)

- structural ambiguity involving lines 9-10

- rightward movement of quantifiers in S_3 (line 7)

- VHD (keyness): 'd (line 3), had (line 10)

---

[41] Readers will note that in the semantics analysis section, the threshold employed for keyness is stricter, at $p < 0.01$. This is because so many items were identified as key that a cut-off point was necessary to ensure the analyses remained workable.

- DD1 (keyness): that (lines 3 and 10), each (line 7)

Internal Deviation

- use of "I" in lines 3, 6

- two fronted time adverbials (lines 5, 10) break previous pattern of placement

- future tense in S_3 (lines 5 to 7)

- perfect aspect in S_2 (lines 2 to 4)

- commissive mood in S_3 (lines 5 to 7)

- obligative mood in S_6 (10 to 11)

- syntactic and typographical patterns do not coincide in line 10

- relative syntactic complexity of S_2 (lines 2-4)

- subject-verb hiatus between lines 3 and 4

- only one hypotactic construction (line 3)

Determinate Deviation

- none

Parallelism

- pattern of main subject deletion in lines 1, 8, 10 (x 2 in line 10)

- appearance of "I" in lines 3, 6

- pattern of possessive ellipsis in lines 4, 11

- N plus postpositive Adj pattern at end of lines 4, 11

- P D N P NP pattern in lines 1, 2

- two time adverbials following verb (lines 1, 3)

- two fronted time adverbials (lines 5, 10)

- one prepositional adverb per stanza (lines 1, 6, 8)

- all sentences in past tense (except for S_3: lines 5 to 7)

- all syntax paratactic (except for S_2: lines 2 to 4)

4.1.3: SEMANTICS

The first technique we can use to identify statistical deviation on the semantic level concerns a concept called 'keyness' (also known as 'distinctiveness'). As Culpeper (2009: 34) writes, "'Keyness' is a matter of [a word] being statistically unusual relative to some norm." To Culpeper's definition, we might add the following by Scott (1997: 236): "A key word may be defined as a *word which occurs with unusual frequency in a given text*. This does not mean high frequency but unusual frequency, by comparison with a reference corpus of some kind." The question of *which* norm or reference corpus can have considerable bearing on the output of the keyness calculation, as well as the validity of the strikingness measure derived from that calculation. However, as an initial justification for the use of this technique, one may point to a crucial notion in the field of text linguistics: this is de Beaugrande and Dressler's (1981: 10-11) concept of "intertextuality" as one of seven defining traits that distinguishes a text from a "non-text" (ibid.: 3). Shreve (2018: 173) provides a useful summary: "When we read a text we do so with a previous experience of texts, and particularly of texts associated with the specific situation in which we are using the text. […] Thus, intertextuality is a reference to the way a given text relates to other texts that are (or have been) relevant in a particular kind of situation or to a specific kind of purpose." Insofar as the keyness calculations performed in this study relate my survey texts to a corpus of similar texts—and the calculations express those relations in terms of statistical deviation, a category of importance to this study—the notion of keyness appears all the more attractive.

However, keyness values are not instant or ready-made predictors of strikingness. There are several other factors that prevent us from using keyness values in that fashion. One is the temporal experience of the text. As Fish ([1970] 1980: 27) has it: "A reader's response to the fifth word in a line or sentence is to a large extent the product of his responses to words one,

two, three, and four."[42] The underlying assumption is that "the reader responds" to the text in terms of its "*temporal* flow" and not to the "whole utterance" instantaneously (ibid.).[43] "That is," he elaborates, "in an utterance of any length, there is a point at which the reader has taken in only the first word, and then the second, and then the third, and so on, and the report of what happens to the reader is always a report of what has happened *to that point*" (ibid.). We can extrapolate from this a hypothetical situation: suppose a key word occurs three times in a text: if strikingness fades with familiarity, and the three instances are not part of some other FG patterning, then there should be some gradation in the strikingness predicted for each instance of that word—a gradation that keyness values cannot provide.

The particular relevance to poetry: while it would be surprising if the effects of FG were limited to poetry, such effects "might be more difficult to observe" with regard to longer works, as Van Peer (1986: 60) writes. The longer a text is, the more likely it is that things other than the text may claim our attention: how many of us could pay complete attention to a single text for more than a handful of hours (without a single unrelated thought intruding)? When it comes to longer texts, non-textual factors may very well influence expectations concerning the occurrence of one word or another. For instance, we might forget that an unusual word has already occurred ten pages before, and when we encounter it again the next time we pick the book up, it is striking to us once again. This seems less likely to occur in a short text. The need for some gradation in our strikingness measures owes in part to the length of the texts investigated here.

While Fish (1973 [1980]: 68-96) could be a fierce critic of stylistics, he nevertheless included in his category of reader response an activity central to much stylistic analysis: "the

---

[42] This assumes that an "aberrant decoding" (see: Eco ([1965] 2003) has not taken place.

[43] While there is no across-the-board agreement on the size of a 'linguistic chunk,' the amount that can be read in a single eye fixation is estimated for "skilled readers of alphabetic writing systems" to span from "three to four letters to the left of fixation (or the beginning of the currently fixated word)" to "fourteen or fifteen letter spaces to the right of fixation" (Rayner, Slattery and Bélanger 2010: 834). This means that even the shortest text in my survey would take more than one eye fixation to read.

projection of […] lexical probabilities [and] their subsequent occurrence or non-occurrence"

(Fish [1970] 1980: 27). I have mentioned expectations informed by intertextuality, but two

other kinds of expectations can also affect the aforementioned "projection" of "lexical

possibilities" (ibid.). These are expectations having to do with collocations and schemata.

Firth (1951: 194) introduced the term collocation in the 1950s, although he only offered

loose definitions (Philip 2011: 39-40). Nevertheless, the general sense can be gleaned from

this description: "The habitual collocations in which words under study appear are quite

simply the word accompaniment, the other word-material in which they are most commonly

or most characteristically embedded" (Firth 1957: 180). While Firth did not specify the length

of text in which these co-occurrences take place, it is generally understood that he was

referring to a "narrow contextual environment" (Scott 1997: 239). Indeed, the narrowness of

this context has been described in terms of the "idiom principle" (Sinclair 1991: 110-115,

173) and the "window of […] collocational force" (Louw 2000: 50).

According to the idiom principle, collocations demonstrate a broader "principle of

organization" in language (Sinclair 1991: 173), whereby a "language user has available […] a

large number of semi-preconstructed phrases that constitute single choices, even though they

might appear to be analysable into segments" (ibid.: 110). Sinclair (ibid.: 172) acknowledges

that the line between idioms and collocations can be blurry, as both involve two or more

words co-occurring in the same short span of text. However, co-occurrences should be

deemed idioms when they provide a "single unit of meaning," such as the phrase "hold sway"

(ibid.). The meaning of this phrase is not necessarily suggested by the individual meanings of

its two words. Collocations, on the other hand, are those phrases in which each word "keeps

some meaning of its own," such as the phrase "hold a meeting" (ibid.). As the preoccupation

with phrases might suggest, the crucial element for Sinclair (ibid.) is proximity, and he relates

that the "usual measure of proximity is a maximum of four words intervening" (ibid.: 170),

223

i.e. four words on either side of the item in question. This is the so-called "nine-word window" that has become something of a convention in corpus research (Bevitori 2015: 113). It is within this window that one is most likely to identify collocates.

As theorised by Louw (2000: 50), this window features in another concept known as "semantic prosody." While the concept can be traced back to Sinclair's (1987) investigation of "the phrasal verb SET IN" (Stewart 2010: 6), the term was popularised by Louw (1993), who defined it as a "consistent aura of meaning with which a form is imbued by its collocates" (ibid.: 157). For example, we may point to the "well-known negative semantic prosody of 'happen'" (Louw and Milojkovic 2016: 296). Because *happen* tends to occur in semantic environments concerned with "unpleasant things—accidents and the like" (Sinclair 1991: 112), it has a negative semantic prosody.[44] The notion of proximity seems crucial in this phenomenon as well. As Louw (2000: 50) writes: "Most semantic prosodies accumulate and concentrate their power within the nine-word window of acknowledged collocational force."

However, collocates are not the only kind of co-occurrence that may influence strikingness. As Scott (1997: 240) argues, there are "socially determined networks of links between ideas" known as "schemata" (see e.g. Schank and Abelson [1977]). As Scott (1997: 233) notes: "The classic example is the restaurant schema, which associates the ideas |menu|, |cook|, |waiter|, |payment|, etc." The problem with focusing on collocations is that items like the above may occur throughout a text but without appearing together within the window of collocational force. As Scott (2000: 111) writes, "[…] relationships between pairs like *dead* and *alive*, *big* and *large*, *fire* and *matches*, *problem* and *solution*," etc., are often not detected, owing to the fact that there is "a low probability of their co-occurring within a 4-word span

---

[44] Or at least that is Sinclair's (ibid.) and Louw's (1993) position. Spencer (2011: 27-28) discusses some studies that problematise this and other cases.

often enough […].” So why not simply increase the collocation span? Scott (ibid.) notes that this is not really an option because “so much noise (unrelated items) enters the listings that the sought-for partner is almost certain to be drowned in the surf of accidentally-proximate items.”

Thus, Scott focuses on what he terms “associates,” or “words which are key in the same text as a given key word” (ibid.: 112). To identify associates, one must investigate the keywords of numerous texts that all have the same key item in common. As it would be prohibitively time-consuming to identify associates of every keyword in the texts analysed here, one must decide which keywords to prioritise in the analysis. This is a crucial choice that concerns the text’s “aboutness” (Scott 2000: 107-109; Phillips 1985). As Scott (2000: 107) has it, this concept concerns “the primary question language users routinely ask of a text.” That question is: “What is it about?” (ibid.). However, I am at risk of putting the cart before the horse, as we have not yet discussed even the initial keyness calculation in sufficient depth.

As with some of the statistical deviation I identified on the phonological and syntactic levels, this technique requires two sets of values to begin with: the observed frequency of an item and its expected frequency. Moreover, we need a statistical test that involves a *p*-value to determine whether the results for a particular lexical item are statistically significant. In the past, this calculation would likely have involved the chi-squared test, but “this has been shown to over-estimate the importance of rare events” (ibid.). For this reason, Dunning (1993) developed the log-likelihood (or G-squared) test in his paper “Accurate Methods for the Statistics of Surprise and Coincidence.”

Since this is a semantic analysis, we are interested chiefly in lexemes (a.k.a. lexical items or semantic words) rather than orthographic or phonological words. That is, we are interested in a “comparatively abstract” unit with a “more or less consistent meaning or function but

which can possibly vary in form for grammatical purposes" (Trask 1992: 158). Trask (ibid.) provides two good examples: "the items *dog* and *dogs* are both particular forms of the lexical item DOG, and *take*, *takes*, *took*, *taking* and *taken* are all particular forms of the lexical item TAKE."

But what about the noun *take* (e.g. 'what's your take on this?)—how should we deal with forms that have the same pronunciation but distinct meanings? This concerns the distinction between homonymy and polysemy. As Croft and Cruise (2004: 111) write, this problem can be viewed either diachronically or synchronically; the "more traditional distinction is the diachronic one," which takes historical etymologies into account. In this view, homonyms are "derived from distinct lexical sources, and their orthographical or phonological identity is due either to the loss of an original distinction due to language change, or to borrowing […]" (ibid.). In other words, homonyms are similar by accident. By contrast, "polysemic units are derived from the same lexical source, being the result of processes of extension such as metaphor or metonymy" (ibid.). According to the *OED*, the noun "take" is derived from the verb; thus, by this definition, it is a polysemic unit. Recourse to etymology illustrates that the diachronic distinction between homonymy and polysemy is a "question of historical fact, resolvable in principle, if not always in practice" (ibid.). But how often do readers have historical etymologies at hand when they read?

It seems much more likely for readers of my survey to make synchronic distinctions; the question here is "whether there is a felt semantic relationship between the two interpretations of a word or not" (ibid.). When it comes to words with the same spelling, Yule (2016: 132) writes that there is a "temptation" to think that they "must be related in meaning." Boase-Beier (2011: 88-89) considers this to be a potential factor influencing translation decisions. She paraphrases Crystal (2003: 191) on this point, writing that "active linguistic context is always potentially separated from collective historical linguistic context" (Boase-Beier 2011:

89). This is another way of saying that similarity on the level of phonology may be enough to motivate the reader to make a "lexical cognitive extension" (ibid.: 88) from one word to another. For these reasons, I will be analysing homonyms as if they were forms of the same lexeme.

The corpus I use first is the poetry subcorpus of the BNC. To compile this subcorpus, I first had to download the entire BNC from the *University of Oxford Text Archive*.[45] Next, I relied on the BNC Web Indexer (Lee 2001) to identify file names containing poetry texts.[46] While Lee's (2001) index referred to an earlier version of the BNC, the file names provided on that site corresponded in all but one case to the files in the BNC XML Edition.[47] Next, I imported the BNC files into AntConc 3.5.7 (Anthony 2018) to create a lemma list. I chose to use the poetry subcorpus of the BNC at this stage—rather than the corpus as a whole— because my surveys present the poems *as poems* (as opposed to another text-type). Drawing on the notion of intertextuality, I hypothesised that readers' previous experiences with poetry might affect their expectations about the lexico-semantic content of the stimulus texts. While internalised knowledge about lexis in the language at large may indeed influence expectations, keyness is better identified using a corpus of texts similar to the ones under investigation (Culpepper 2009: 34-35). Thus, at this stage both the target corpus and the reference corpus should consist of poetry. To calculate keyness against the entire BNC, or against some other written subcorpus thereof, would be counterproductive for our purposes. The point is not to flag up features that might be striking to the reader on the basis of differences between poetry and, say, informative writing. The point is to flag up features that would stand out against the backdrop of readers' experiences with the text-type in question.

---

[45] The version downloaded was the XML Edition (2007).

[46] This indexer can be found at: http://ucrel.lancs.ac.uk/bncindex/.

[47] The file in question was labelled "CHX." This work appears not to have been included in the BNC XML Edition, perhaps owing to copyright issues. Such deletions have happened before (Leech, Rayson and Wilson 2001: 4.)

This picture is complicated by the fact that (apart from in Survey 3), the poems are presented not just as poems, but as translations from the Japanese, and respondents *may not bring the same lexico-semantic expectations to Japanese poems as they would to English ones*. However, since there is no corpus of translated poems readily available to me for analysis, a corpus of texts originally written in English may well be the only other option where the procedure is concerned.

There are two things to note about my keyness analyses. First: for words with multiple grammatical functions, the choice arises whether to seek out the BNC frequencies associated with specific grammatical functions, or to use the frequencies associated with their lemmas. To rephrase this, the choice is between using a smaller or larger frequency—that of the variant or the lemma. Throughout my analyses, I have used the larger frequency. The reason is: while we are mostly concerned with the positive keyness values associated with content words, negative keyness values are also possible. These would indicate words occurring at an unusually low frequency. Theoretically, the sudden appearance of an item that has hitherto been underused might stand out. Of course, it is impossible to have respondents indicate this kind of strikingness without the item occurring at least once. Nevertheless, using the higher frequency value (the lemma value) would mathematically give negative keyness more of a chance to emerge.

The second thing to note concerns hyphenated words. I have treated "half-open" in Beichman's "Shijimi Clams" as two separate lexemes, but have treated "light-years" and "light years" (in Elliott and Kawamura's text and Bownas and Thwaite's text, respectively) as single lexemes. This is consistent with Crystal's (2008: 276) definition of lexemes as "the units which are conventionally listed in dictionaries as separate entries."[48] This also concerns the notion of rarity discussed earlier in connection with the log-likelihood method. According

---

[48] That is, "light-year" or "light year" are more likely to have their own dictionary entries than "half-open."

to the BYU-BNC, the string "half-open" appears only 73 times in the entire BNC corpus, suggesting the word is very rare. But is a word like "half-open" really just this side of "apoptosis" (66 occurrences according to the same search) or "yanomami" (with 67 occurrences) in terms of unexpectedness? It is likely that expectations concerning this word are more accurately modelled using frequencies for its individual components.

Fortunately, there is no need to perform log likelihood calculations by hand. Instead, we may use a programme like AntConc (Anthony 2018)—although this will come with a caveat or two. AntConc's Keyword List function yields the following lemmatised keyness values for Beichman's "Shijimi Clams":

Fig. 20. Keyness Values

Rank	Freq	Pos/Neg	LL	Word
1	2	+	32.96	clam
2	1	+	16.46	cackle
3	1	+	16.46	shijimi
4	2	+	13.03	mouth
5	2	+	12.13	open
6	1	+	11.97	gobble
7	1	+	11.06	witch
8	2	+	10.35	night
9	1	+	8.75	evil
10	1	+	8.08	alive
11	1	+	7.58	couldn't
12	1	+	7.04	buy
13	1	+	6.99	help
14	1	+	6.99	kitchen
15	1	+	6.57	wake
16	1	+	6.53	dawn
17	1	+	6.29	corner
18	2	+	6.17	up
19	1	+	5.91	evening
20	1	+	5.8	half
21	1	+	5.75	i'll
22	1	+	5.24	i'd
23	1	+	4.87	let
24	1	+	4.81	sleep
25	1	+	4.69	dead
26	1	+	4.21	every
27	1	+	4.16	little
28	1	+	3.83	after
29	2	+	3.48	that
30	1	+	3.34	each
31	1	+	3.3	old
32	1	+	2.31	were
33	1	+	2.2	had
34	1	+	1.8	one
35	1	+	1.63	out
36	1	+	1.52	all
37	1	+	1.27	an
38	1	+	1.03	at
39	1	+	0.79	like
40	2	+	0.67	in
41	1	+	0.53	you
42	1	+	0.39	with
43	1	+	0.31	it
44	2	+	0.11	of
45	2	+	0.03	a

Now for the caveats I mentioned. Most forms have been properly lemmatised—for instance, when we look at item 4 (*mouth*), we see from the frequency column that two occurrences are accounted for (these are 'mouths' in line 4 of the poem, and 'mouth' in line 11). However, *a* (item 45) and *an* (item 37) are not treated as variants of the same lemma. Moreover, several contracted forms appear in this list: item 11 (*couldn't*), item 21 (*I'll*) and item 22 (*I'd*). In the case of item 22, the *'d* in *I'd* would ideally have been lemmatised together with item 33 (*had*). This is a function of the particular lemma lists employed, and of

the way that AntConc identifies tokens.

Nevertheless, for the vast majority of items, the keyness values (in the LL column) are very similar to ones I calculated using the BYU-BNC and Rayson's (2016) "Log-Likelihood and Effect Size Calculator" in a preliminary attempt. In fact, the same ten items appear as the top ten keywords in both lists (with only *gobble* and *open* transposed). Given that the differences in lemmatisation are confined to a handful of function words, I believe the use of AntConc in this project can be justified on the basis that it saves a considerable amount of time in an already time-intensive process.

How is the above list useful to us? I have already mentioned keyness values as a potential, albeit rudimentary, predictor of strikingness, insofar as they describe frequencies that would be unusual for texts of a particular type. However, as I discussed earlier in this section, to use these values alone would be to ignore the effects of 'aboutness' and schemata.

As Yablo (2014: 1) notes, the study of aboutness is nothing new; it can be traced back to Brentano and Husserl, who were themselves informed by medieval thought on intentionality. I should note, this is not the same intentionality discussed by Wimsatt and Beardsley (1946). The idea is rather that of "mental directedness towards (or attending to) objects" (Jacob 2019; n.p.). As Jacob (ibid.) writes, this concept concerns "the power of minds and mental states to be about, to represent, or to stand for, things, properties and states of affairs. To say of an individual's mental states that they have intentionality is to say that they are mental representations or that they have contents." In short, aboutness is "the relation that meaningful items bear to whatever it is they are *on* or *of* or that they *address* or *concern*" (Yablo 2014: 1).

In the field of linguistics, Phillips (1985: 3) used the term to investigate what it is that "we mean when we speak of the 'subject matter' of such and such a book or the 'topic' of chapter so and so." The problem he was interested in was how to explain the mechanism(s)

underlying the perception of aboutness. His hypothesis was that aboutness "stems in part from the reader's appreciation of certain global patterns of textual organisation" which he called macrostructures (ibid.: 4). To test his hypothesis, he performed a cluster analysis on a selection of science textbooks and novels. As evidence of macrostructure, he decided to focus on "patterns of association" (ibid. 100) between lexical units within each text. As for why he decided to focus on lexis, the assumption was that the choice of a lexeme is an "aspect of authorial control which results from the choice of topic" (ibid.: 61). His findings revealed strong evidence of inter-chapter similarity for the science texts (ibid.: 195-196). By contrast, for the non-science texts, there was "very little similarity of networks across text intervals and as a result no indication of significant chapter linkages" (ibid.: 202).

For Phillips (ibid.: 224), the most plausible explanation for these results "lies in the nature of the reality which is projected" by the non-science texts as opposed to the science texts. There are "additional layers of symbolism which distance a literary work further from experienced, phenomenological reality," and this accounts for the "inability to detect macrostructure in the non-science texts" (ibid.: 226). In other words, to identify the aboutness of a science text, one can presumably point directly to items in the text itself; however, this procedure is unlikely to yield the aboutness of a literary text.

If it seems like we are dealing with two types of aboutness here, the suspicion is not unfounded. Scott (2000: 109), for instance, makes a distinction between text-dependent and reader-dependent aboutness: "Orwell's *Animal Farm* is undeniably 'about' farmyard animals, but only 'arguably' about totalitarianism. So aboutness can refer either to explicit signals in the text, or to underlying, implicit significance." By the same token, we can also discuss the science texts in Phillips' (1985) study in terms of either kind of aboutness. For instance, it could be argued that science textbooks are in some way about humanity's desire to better understand and/or control the natural environment.

While Phillips (1985: 216) reports that "the non-science texts show little tendency to macrostuctural organisation," a keyword analysis would presumably have flagged up elements connected to explicit aboutness. Moreover, as I argue below, in the context of poetry, explicit aboutness as signalled by titles would seem likely to influence expectations concerning lexical probabilities. As will we see, this will involve the notion of schemata touched on earlier in this section.

Firstly, as Scott (2000: 110) relates, the question of how aboutness is "flagged up for the reader or listener" has many potential answers. It could be flagged up in the title, concluding remarks, in repetition, etc (ibid.). Given that I will account for aboutness in these FG analyses by compiling lists of "associates" (Scott 1997, Scott 2000)—and given that it will not be possible to investigate associates for every keyword in the poems—it is imperative to choose a reliable indicator. Insofar as title words in each of the titled poems in my target corpus reappear in the body of the poems, the titles seem a likely candidate. Further support for this decision comes from the empirical studies discussed below.

There have been numerous studies of how headings affect text processing at the level of recall and/or comprehension (see e.g. Polley Sanchez, Pugzles Lorch and Lorch [2001]; Lemarié, Lorch and Péry-Woodley [2012]). As Lemarié, Lorch and Péry-Woodley (2012: 6) write, previous research has established that headings "influence memory and comprehension of text" in at least 3 different ways:

> First, when headings are used to highlight the organization of topics in a well-structured text, they lead to better memory for that organization […]. Second, titles emphasize specific topics or themes, which biases readers' understanding of the text in the direction of the emphasized topics and themes. Finally, by establishing a context, headings can influence the interpretation of text content by causing readers to use relevant background knowledge to guide

233

comprehension.

The second and third mechanisms in the above list have been understood in terms of the activation of schemata (see Brooks et al. 1983 and Wiley and Rayner 2000). In the reference above to topics and themes, we see an overlap with the concept of aboutness as conceptualised by Phillips (1985) and Scott (2000): the notion of the topic of a text seems congruent with our discussion of explicit aboutness, while the idea of the theme would not be a bad way to explain implicit aboutness. So if titles may be considered a good indicator of aboutness (insofar as they activate schemata), and if associates can help us identify those schemata as Scott (1997) suggests, then it makes sense to try to identify associates of title words. Of course, a reader of poetry may well be able to cite numerous poems whose titles have little or no immediate connection to the body of the poem, whereas such a titling practice would presumably be less common in other text-types. However, this does not seem to apply to the poems in my survey. The titles repeat items used in the main bodies of the poems (and key items, at that).

Thus, in the next part of this analysis, we will see whether any schema items (as determined by analysis of associates) turn up subsequently in the body of the poem. If these are part of the schemata activated at the outset of the poem, it is not unreasonable to suppose that these semantic concepts might be less striking to survey respondents, and we can adjust our strikingness predictions accordingly.

For the purposes of this analysis, we can identify two kinds of schemata: frames (Minsky 1975) and scripts (Schank and Abelson 1977). The former represent knowledge about objects while the latter include knowledge about actions and events. They both involve a theoretical construct called "slots," which are occupied by entities called "fillers" (Taboada 2004: 20). Thus, the frame for the concept HOUSE will have "slots with labels such as 'living room,' 'bedroom,' 'kitchen,' etc." (ibid.). On the other hand, scripts have "not only objects, but also

typical actions as slot labels" (ibid.). Thus, a script for the concept RESTAURANT will include "'table,' 'food,' 'menu,' 'waiter/waitress,' but also 'order,' 'eat,' 'pay,' 'tip,' as some of its empty slots, with some optionality available, depending on the type of restaurant […]" (ibid.).

To identify associates for Beichman's "Shijimi Clams," I used the online corpus tool Sketch Engine (Kilgarriff et al., 2004) to compile a corpus of webpages returned from a search of the words *clam*, *clams*, *clamming*, and *clammed*. I set the Max URLs option to 100 (though only 80 results were returned, yielding a corpus of 67,299 tokens as identified by AntConc). Of these, I chose the first ten files (totalling 10,439 tokens) and created a .txt file for each. Since batch-processing was not possible in AntConc, I loaded these files into a different piece of software called WordSmith Tools 4.0 (Scott 2004) to determine keywords for each file. As my reference corpus, I decided to use the BNC word list offered by the software's creator at:

https://lexically.net/downloads/BNC_wordlists/downloading%20BNC.htm.

The entire BNC was used here, as opposed to a specific subcorpus, on the assumption that since the target corpus (the webpages) embody a range of text-types, the reference corpus should be representative of language use in general. In compiling the web corpora for these analyses, it was not possible to specify the domain endings (e.g. .com or .co.uk) of search results, so the corpora reflect a mix of sources. Thus, one possible objection to using the BNC as a reference corpus is that this risks flagging up items that are more common in American English, since the frequencies of such items (i.e. eggplant as opposed to aubergine) would presumably be low in the BNC.

However, in actual practice, this does not seem to have happened. The keyness results I obtained using the BNC as my reference corpus were very similar to those I subsequently obtained using a different reference corpus. (For that reference corpus, I used my entire

Sketch Engine corpus, a corpus of over a million tokens compiled from search results revolving around *clams*, *shellfish*, *isolation*, *solitude*, *light-years*, *love* and *rain*).[49] This similarity, both in the number of keywords and in the specific keywords themselves, is rather surprising. After all, the Sketch Engine corpus contains many .com results, and thus probably contains a lot more American English than the BNC. Furthermore, it consists only of texts related to the above-named concepts, and we know that those concepts represent a rather small subset of potential topics in English. One would not expect it to be a very balanced corpus. Nevertheless, these initial results recall a passage in Scott (2000: 115) wherein he discusses identifying keywords in an academic text by using a reference corpus of *Guardian* newspaper texts: "No claim is made (or could be) that this sort of academic text is similar to Guardian [sic] journalism, although in practice I do not find that changing the reference corpus makes much difference." This claim would seem to run contrary to the principles of corpus selection discussed earlier by Culpepper (2009).

The resulting keyword lists for the target corpora (the first ten files from Sketch Engine related to clams) contain occasional lemmatisation errors (attributable not to the programmes, but to the specific lemma lists used).[50] Some words are identified as key in multiple texts. These are what Scott (1997) calls 'key keywords,' or "*words which are key in a large number of texts of a given type*" (ibid.: 237). Ten texts would not normally count as a large number, but it was the largest *feasible* number to work with in this project, given the time it takes to prepare each file. In the key keywords data below, the third column shows how many texts

---

[49] The first five items are title words from the titled texts. "Love" and "rain" were chosen from the untitled texts (the *Midaregami* translations). Because those poems were untitled, I needed to identify some other words that could be hypothesised to guide the interpretation of the poem. In the end, it struck me that I had overlooked cultural salience: was there any word of particular social or cultural importance that could guide readers' interpretation? It turned out there *was* one concept of immense cultural import that "appears to be a nearly universal phenomenon, appearing in every culture for which data are available" (Aron, Fisher and Strong 2006: 595). That is: love.

[50] For this analysis, I set the p value at $p < 0.05$ and set the minimum frequency value at 2. The justification for this value can be found in Scott (1997: 244n9). For my lemma list, I used "Lemma List 10" from: https://lexically.net/wordsmith/support/lemma_lists.html.

(out of the ten analysed) a particular keyword has appeared in. Thus, item 6 (*sand*) was

identified as key in 7 texts.

Fig. 21. Key Keywords List: Associates of *Clam, Clams, Clamming, Clammed*

N	KW	Texts	N	KW	Texts	N	KW	Texts	N	KW	Texts
1	CLAM	10	21	S	4	41	HOLE	3	61	ALTHOUGH	2
2	CLAMS	10	22	SHELL	4	42	INCHES	3	62	ANY	2
3	DIG	7	23	SPECIES	4	43	LIMIT	3	63	AREA	2
4	DIGGING	7	24	#	3	44	LONG	3	64	AROUND	2
5	OR	7	25	ABOUT	3	45	NEAR	3	65	BEACHES	2
6	SAND	7	26	AREAS	3	46	NORTH	3	66	BURROWING	2
7	CLAMMING	6	27	AVOID	3	47	OREGON	3	67	BUTTER	2
8	ARE	5	28	BEACH	3	48	SHELLFISH	3	68	CAREFUL	2
9	BAY	5	29	BOAT	3	49	SHELLS	3	69	CHOWDER	2
10	HARVEST	5	30	CALLED	3	50	SHOVEL	3	70	CLOSED	2
11	LOW	5	31	COMMONLY	3	51	SHOW	3	71	COAST	2
12	MUD	5	32	DEEP	3	52	SOFT	3	72	COASTAL	2
13	OCEAN	5	33	DIGGERS	3	53	SURE	3	73	COMMERCIAL	2
14	TIDE	5	34	EATEN	3	54	UP	3	74	COMMON	2
15	BROKEN	4	35	FEW	3	55	USE	3	75	CONSERVATION	2
16	CAN	4	36	FIND	3	56	YOU	3	76	CONTACT	2
17	DUG	4	37	FISHING	3	57	YOUR	3	77	CUT	2
18	FOOT	4	38	FOUND	3	58	A	2	78	DEPARTMENT	2
19	HARD	4	39	HAND	3	59	ALONG	2	79	DIGGER	2
20	RAZOR	4	40	HARVESTED	3	60	ALSO	2	80	DOESN	2

N	KW	Texts	N	KW	Texts	N	KW	Texts	N	KW	Texts
81	DON	2	101	LARGE	2	121	QUAHOG	2	141	USING	2
82	DUE	2	102	LENGTH	2	122	RECREATIONAL	2	142	WATER	2
83	DURING	2	103	LICENSE	2	123	REGARDING	2	143	WE	2
84	EASY	2	104	LIKE	2	124	SEA	2	144	WHEN	2
85	EDIBLE	2	105	LIVE	2	125	SHARP	2	145	WINTER	2
86	EQUAL	2	106	LIVING	2	126	SIZE	2	146	YEARS	2
87	ESPECIALLY	2	107	MAINE	2	127	SMALL	2	147	YOU'VE	2
88	FALL	2	108	MANY	2	128	SOUTH	2			
89	FAMILY	2	109	MOLLUSKS	2	129	STATES	2			
90	FISH	2	110	MOVE	2	130	STRAIGHT	2			
91	FIVE	2	111	MUSSELS	2	131	T	2			
92	FLATS	2	112	NECK	2	132	TASTE	2			
93	FROM	2	113	ODFW	2	133	THAT	2			
94	FUN	2	114	OFTEN	2	134	TIDES	2			
95	HALF	2	115	OTHER	2	135	TILLAMOOK	2			
96	HANDLE	2	116	OUR	2	136	TIPS	2			
97	HIGH	2	117	PACIFIC	2	137	TO	2			
98	IS	2	118	PLACE	2	138	TOXINS	2			
99	KEEP	2	119	POPULATION	2	139	TRY	2			
100	KNOWN	2	120	PROBABLY	2	140	UNITED	2			

As might be expected, we find numerous items related to food schemata (34: eaten; 67:

butter; 69: chowder; 85: edible; 132: taste). Judging by these associates, some food schemata

are probably activated by the title. Let us think back to the keyness list from the poem. If we

consider as statistical deviation only those items that are significant at $p < 0.01$ (this

corresponds to a keyness value of 6.63 or higher), then we may safely demote the words

"mouth", "kitchen," and "gobble." Although these do not appear on our key keywords list,

the surprise of encountering them is probably diminished by their belonging to schemata activated by the title. "Mouth," however may still stand out as part of a parallelism, as lines 4 and 11 establish a likeness between speaker and clam. We can add this to our FG list. Moreover, this likeness seems to suggest the common conceptual metaphor YOU ARE WHAT YOU EAT, but at the same time this metaphor is subverted insofar as the poem re-figures the eater as a kind of malevolent captor. We could arguably classify this as an instance of determinate deviation, only it is unclear precisely what part of the poem a reader would underline to indicate this. Similarly, "gobble" may still stand out as it seems to mark a shift in register. We can add this under the heading of Internal Deviation. Moreover, our key keywords list contains the items *live* and *living*. These occur in two websites each. (In the case of one website, *living* refers not to clams, but to people. However, if the concepts *live* and *living* were grouped together semantically, the other website with *living* in it could be grouped together with the two that contain the verb *live* in reference to clams.) The idea that clams are living creatures would probably be a part of most readers' frame for the concept CLAM, so we may demote the word "alive" in line 4 of Beichman's translation. However, this word may still stand out insofar as it involves a mild internal deviation—the first reference to deadness or aliveness in the poem is to time ("the dead of night"), while the second is to an organism. No other content word in Beichman's "Shijimi Clams" appears as an associate of *clam* on our key keyword list.

It would have been convenient if *mouth* and *open* had appeared in our key keywords list, or even in the list of collocations returned by the BYU-BNC website for a search of [clam]. It seems likely that these items would be part of most readers' frames for the concept CLAM. That is, it is probably not a stretch to assume that most readers surveyed would describe a clam as something with a mouth that can open and close. As it stands, I have demoted *mouth* on the basis that CLAM can be seen not only as its own schema, but as part of other

schemata—including the schema for FOOD. That is, it is common knowledge that clams are a kind of food. FOOD, alongside other items like MOUTH, would presumably be part of the script for EATING. This raises an interesting question with methodological ramifications: how far up or down these schematic chains should the researcher be permitted to travel in search of items to cross off the keywords list (i.e. items that might be expected to co-occur with *clam*)? Thus, we see the issue of subjectivity creep into what seemed at first like an objective procedure. This puts the researcher in an unenviable position: does one base all decisions strictly on the key keywords list and thereby open oneself to the charge of ignoring common sense (where common sense says *mouth* and *open* are associated with the concept CLAM), or does one use common sense and open oneself to the charge of making decisions that the keywords data does not strictly support? At present, the only way out of this quandary that I see is to re-examine the local context of problematic items as they arise. The fact that *open* appears right after *mouth* in the poem—that is, within the window of collocational force—suggests that we may demote *open* on the basis that it is a common collocation of *mouth*. (A search for collocations of [mouth] in the BYU-BNC bears this out.) Ideally, one would be able to identify and cross-reference collocations and associates of all items in the poem. However, since that would be enormously time-intensive, I will only be using this additional procedure when strictly necessary. ***At the end of the day, we are talking about two different—but theoretically interrelated—kinds of strikingness: one is strikingness based on intertextual comparison (which is what the keyness values indicate), and the other is strikingness based on text-internal mechanisms.*** There is nothing in the literature on FG to suggest which one of these should take precedent, especially when it comes to texts as short as the ones I am investigating. However, it does not seem unreasonable to suppose that this brevity would give readers less time to adjust to the poem as its own text. Therefore, the surprise of encountering statistically unusual words might

trump some of the effects of collocation. The assumption is that collocations of words like *mouth*—i.e. a word that is already associated with CLAM—will be less striking than collocations of words like *witch*, which among our keywords has perhaps the least to do with the concept CLAM. Thus, I have demoted *open* while letting *evil* stand. It is hoped that the strategy employed here strikes an appropriate balance when it comes to accounting for both intertextual and text-internal strikingness.

The key keyword list suggests a somewhat wider variety of concepts associated with clams than the list of collocations returned by the BYU-BNC website for a search of [clam]. For instance, present in the key keyword list but missing from the collocation data are *sand*, *harvest*, *mud*, *breaking*, *holes*, *shellfish*, *shovels*. However, whether our key keyword list actually represents an increase in the number of schemata over the BYU-BNC collocation data is arguable. For instance, while the item *sand* (flagged up as a key keyword) does not occur in the collocation data, the word *beach* does—and *sand* is likely to fill a slot in most readers' frames for the concept BEACH. The question of whether our key keyword list indicates more schemata than the collocation data also depends on whether we impose a minimum frequency on the collocations. The word *beach*, for instance, only occurs once on the collocation list, and it is not even in the top 100 results.

Nevertheless, the collocations list is not without its uses. For instance, it reveals the word *up* (as in 'clam up') to be the most frequent collocate of [clam], reminding us not to forget the influence that idioms and dead metaphors may exert on word-choice expectations. For instance, there is also the expression "happy as a clam" in English. Kovecses (2010: 99) identifies this as an example of the conceptual metaphor A HAPPY PERSON IS AN ANIMAL (THAT LIVES WELL). Indeed, it is not unreasonable to suppose that 'clamming up' and 'happy as a clam' might be part of the background knowledge activated by the title, thus influencing expectations concerning lexis. While neither of those phrases actually

240

appears in the poem, there *are* two references to mouths being open (the opposite of clamming up). But am I really justified in adding these to my FG list as subversions of the 'clamming up' metaphor? Are they likely to stand out for that reason? Unfortunately, the previous literature does not provide any guidance in cases like this. I would be more inclined to add them if the local context had more items related to talking rather than eating.

There are a number of other semantic FG features to identify in Beichman's translation. The poem opens with what could be described as a mild schema violation involving the script for WAKING UP. Given that in the present day it is not uncommon for people to work night jobs, nor is it uncommon for people to get up in the middle of the night for any number of reasons, we may not be able to posit a very strong association between waking and the morning hours. However, collocation data for [wake] from the BYU-BNC still suggests a stronger association with "morning" (340 hits) than with "night" (152 hits). Similarly, collocation data for the exact string "woke up" yields 87 hits for "morning" and 22 hits for "night." Thus it enters our FG inventory as a determinate deviation.

Next, there is some ambiguity in the inferential links between sentences. For instance, does the second sentence ("in a corner of the kitchen the little clams I'd bought […] were alive") imply that the speaker has gone to the kitchen after waking? Or is she simply musing in bed? Similarly, it is unclear whether the material in stanza 2 is actually spoken aloud to the clams or is simply a representation of the speaker's thoughts at that moment. As Saeed (2003: 203) writes, speakers "seem confident" that listeners will make "inferences to preserve a notion of coherence in what they are told"; furthermore, speakers "take advantage of [this] to speak less explicitly than they might." However, this tendency is not to be confused for a natural propensity toward ambiguity; on the contrary, Grice's (1975) maxims describe the opposite: listeners usually assume that "a speaker will have calculated her utterance along a number of parameters" (Saeed 2003: 205), one of which is 'perspicuity.' That is, there is an

expectation that a speaker will avoid both ambiguity and obscurity, and will communicate in a brief and orderly fashion (Grice 1975: 46). Thus, when speakers take advantage of inference-making on the part of listeners, they are fulfilling the remit of being brief. However, while the sentences above would seem to invite inference-making on the part of readers, they do not avoid ambiguity. Insofar as they contravene Grice's maxim of manner, we may enter them into our FG list as determinate deviations. (It is supposed here that such ambiguity would make sentences 2 and 3 more likely to stand out in their entirety.)

Stanza 2 subsequently suggests a schema violation concerning the interaction between humans and clams: most readers' scripts for dealing with clams are unlikely to include talking to them. It may well be the case that such behaviour is expected in the context of a fairy tale, but until the change of register in stanza 2, there are no clues that such a context would be relevant. We may enter this as an instance of determinate deviation.

The change of register which positions the speaker as a kind of captor or fairy tale character may well prepare readers to accept a comparison like the one that subsequently appears in line 9, wherein the speaker compares herself to a witch. However, we need not demote the item *witch* (previously identified as a keyword) as a candidate for strikingness— the change in register may very well activate knowledge related to fairy tales, but it seems unlikely that a reader would be able to predict which particular fairy tale items would appear in the third stanza. (The reader may indeed guess that *witch* would appear after *cackle*, but accounting for the collocational behaviour of every word in the poem in this fashion is simply unfeasible). While the idea of cackling may be consistent with the behaviour of a witch, on another level it would seem to violate the conventional script for EATING. We may add this as an instance of determinate deviation.

The translation also establishes a pattern of backward-looking (i.e. anaphoric) devices of cohesion in phrases like: "I'll gobble you up" (where the "you" refers back to the clams

mentioned earlier in the text), "each and every one" (where the "one" points back to "you"), and "after that" (where "that" refers back to the speaker cackling like a witch). This pattern is broken in lines 10 and 11, where the speaker says "couldn't help it, had to / sleep all night with mouth half-open." In this case, the word "it" refers *forward* to the action of sleeping all night. We may classify this instance of cataphora as an internal deviation.

There is an unusual implicature in the third stanza concerning the last two sentences. That is, after deciding to eat the clams in her kitchen at dawn, the speaker seems to indicate that only one course of action could be expected: to sleep with her mouth half-open. This seems to carry with it a set of assumptions unlikely to be shared by readers, perhaps leading them to conclude that the world in the poem is not really *the* world but merely some "possible world" (see Van Peer 1980: 188). Even if readers merely assume that the speaker's perception of *the* world is idiosyncratic, this would still mark a break with the rest of the poem, at least in terms of implicature; in our previous discussion of inferential links between sentences, we merely noted that multiple inferences could be drawn, leading to a condition of ambiguity. In this case, however, it is the very basis for the inference itself that is unlikely to be shared by readers. We may classify this as an instance of determinate deviation insofar as it seems to depart from conventions external to the text.

Finally, we might point to two features of the last line of the poem. In the concept "half-open" there seems to be a break in a pattern involving the completeness of actions. That is, in line 4 the clams' mouths are open (the action is complete). Moreover, the phrase "gobble you up / each and every one" would suggest an eating action that is also complete. The same goes for "let out" in "let out a cackle" (line 8). By contrast, the speaker's mouth is only half-open in line 11. Thus, we may classify "half-open" in line 11 as an instance of internal deviation. But there also seems to be a kind of logical contradiction in the last line of the poem. The phrase "all night" presupposes that the speaker still has the entirety of the night available to

sleep. However, as we know, the speaker has woken up in the "dead of night," so that some of the night has already passed. This is an example of what is called presupposition failure (see Yablo 2006: 164). We can add this to our list under the heading of Determinate Deviation.

Statistical Deviation

- "clam" in title (keyness)

- "cackle" in line 8 (keyness)

- "Shijimi" in title (keyness)

- "witch" in line 9 (keyness)

- "night" in line 1 (keyness)

- "evil" in line 9 (keyness)

- "couldn't" in line 10 (keyness)

- "bought" in line 3 (keyness)

- "help" in line 10 (keyness)

Internal Deviation

- shift in register signalled by "gobble" in line 6

- shift of deadness/aliveness from metaphorical to literal category ("alive" in line 4)

- break in pattern of completeness of actions (line 11)

- cataphora in lines 10, 11 (break in pattern of anaphoric reference)

Determinate Deviation

- schema violation ("woke up in the dead of night") in line 1

- ambiguity in inferential links (between first and second sentence, and second and third sentence) (lines 1-2 & 4-5)

- schema violation (talking to clams) in stanza 2 (lines 5-7)

244

- schema violation (cackling as part of EATING schema) in stanza 3 (line 8)

- presupposition failure ("sleep all night") in line 11

- unusual implicature (lines 10-11)

Parallelism

- clams and speaker described in terms of mouth/openness (lines 4, 11)

- repetition of "clam" (title and line 3)

4.2. LINES RANKED

4.2.1. BEICHMAN'S "SHIJIMI CLAMS"

The table below tallies all items from the preceding FG lists for Beichman's "Shijimi Clams." The Phonology column displays how many FG features per line we entered into our list of FG on the phonological level. The Syntax column displays the same for our analysis of FG on the syntactic level, and the Semantics column displays the same for our FG analysis on that level.

Fig. 22. Table of FG (Poem A)

	Phonology	Syntax	Semantics	Total
Title	8	1	3	12
Line 1	6	7	3	16
Line 2	9	4	1	14
Line 3	12	10	2	24
Line 4	9	8	3	20
Line 5	6	5	2	13
Line 6	10	7	2	19
Line 7	11	5	1	17
Line 8	6	4	2	12
Line 9	5	1	2	8
Line 10	8	11	4	23
Line 11	7	4	5	16
Total	97	67	30	

Van Peer's (1986) assumption is that lines with higher FG totals are more likely to be underlined, in whole or in part, by readers. To test this assumption, we can use the values from the Total column in the above table to assign a rank to each verse line. This ranking represents how likely we think it is that each line will attract underlining, from most likely (1) to least likely (12). This ranking can be seen in the table below.

Fig. 23. Ranking of Lines (Poem A)

Verse Line	Total FG Features	Rank
Line 3	24	1
Line 10	23	2
Line 4	20	3
Line 6	19	4
Line 7	17	5
Line 1	16	6
Line 11	16	7
Line 2	14	8
Line 5	13	9
Title	12	10
Line 8	12	11
Line 9	8	12

Using the table above, we can now assign each line to one of two categories. The first category consists of those lines more likely to attract underlinings. We will call it FG. The second category consists of those lines less likely to attract underlinings. We will call it BG. The division of the lines into two groups can be seen below:

Fig. 24. Division of Lines (Poem A)

Verse Line	Rank	FG/BG			FG	BG
Line 3	1	FG		No. of lines	6	6
Line 10	2	FG				
Line 4	3	FG				
Line 6	4	FG				
Line 7	5	FG				
Line 1	6	FG				
Line 11	7	BG				
Line 2	8	BG				
Line 5	9	BG				
Title	10	BG				
Line 8	11	BG				
Line 9	12	BG				

Since Beichman's "Shijimi Clams" consists of an even number of lines (including the title), there are an equal number of lines in the FG and BG categories. In cases where there are an odd number of lines (including title), a choice must be made about the centrally ranked line: which category do we put it in? Following Van Peer (1986: 97), I have assigned such lines to the BG category. This choice is *unfavourable* to the theory of FG and the hypotheses derived from it. In other words, this raises the bar for the evidence that must be presented to support the theory of FG.

4.2.2. REXROTH AND ATSUMI'S "SHELLFISH"

The tables below show all relevant FG totals and line rankings for this text.

Fig. 25. Table of FG (Poem E)

	Phonology	Syntax	Semantics	Total
Title	4	1	2	7
Line 1	6	5	6	17
Line 2	15	10	5	30
Line 3	8	6	1	15
Line 4	14	11	4	29
Line 5	11	2	6	19
Line 6	14	5	4	23
Line 7	13	6	6	25
Total	85	46	34	

Fig. 26. Ranking of Lines (Poem E)

Verse Line	Total FG Features	Rank
Line 2	30	1
Line 4	29	2
Line 7	25	3
Line 6	23	4
Line 5	19	5
Line 1	17	6
Line 3	15	7
Title	7	8

Fig. 27. Division of Lines (Poem E)

Verse Line	Rank	FG/BG			FG	BG
Line 2	1	FG		No. of lines	4	4
Line 4	2	FG				
Line 7	3	FG				
Line 6	4	FG				
Line 5	5	BG				
Line 1	6	BG				
Line 3	7	BG				
Title	8	BG				

4.2.3. ELLIOTT AND KAWAMURA'S TWO BILLION LIGHT-YEARS OF SOLITUDE

The tables below show all relevant FG totals and line rankings for this text.

Fig. 28. Table of FG (Poem F)

	Phonology	Syntax	Semantics	Total
Title	6	1	4	11
Line 1	6	1	4	11
Line 2	2	2	4	8
Line 3	4	2	3	9
Line 4	4	2	4	10
Line 5	7	2	7	16
Line 6	13	7	5	25
Line 7	4	2	3	9
Line 8	1	2	1	4
Line 9	6	1	4	11
Line 10	3	3	2	8
Line 11	4	3	6	13
Line 12	6	3	6	15
Line 13	5	4	5	14
Line 14	4	3	6	13
Line 15	10	5	8	23
Line 16	6	3	5	14
Total	91	46	77	

Fig. 29. Ranking of Lines (Poem F)

Verse Line	Total FG Features	Rank
Line 6	25	1
Line 15	23	2
Line 5	16	3
Line 12	15	4
Line 13	14	5
Line 16	14	6
Line 11	13	7
Line 14	13	8
Title	11	9
Line 1	11	10
Line 9	11	11
Line 4	10	12
Line 3	9	13
Line 7	9	14
Line 2	8	15
Line 10	8	16
Line 8	4	17

Fig. 30. Division of Lines (Poem F)

Verse Line	Rank	FG/BG			FG	BG
Line 6	1	FG		No. of lines	8	9
Line 15	2	FG				
Line 5	3	FG				
Line 12	4	FG				
Line 13	5	FG				
Line 16	6	FG				
Line 11	7	FG				
Line 14	8	FG				
Title	9	BG				
Line 1	10	BG				
Line 9	11	BG				
Line 4	12	BG				
Line 3	13	BG				
Line 7	14	BG				
Line 2	15	BG				
Line 10	16	BG				
Line 8	17	BG				

4.2.4. BOWNAS AND THWAITE'S "THE ISOLATION OF TWO MILLIARD LIGHT

YEARS"

The tables below show all relevant FG totals and line rankings for this text.

Fig. 31. Table of FG (Poem B)

	Phonology	Syntax	Semantics	Total
Title	7	1	6	14
Line 1	5	4	4	13
Line 2	4	8	2	14
Line 3	5	3	3	11
Line 4	3	5	5	13
Line 5	8	5	3	16
Line 6	8	3	7	18
Line 7	10	3	4	17
Line 8	5	3	1	9
Line 9	11	2	3	16
Line 10	7	3	4	14
Line 11	9	5	5	19
Line 12	7	4	6	17
Line 13	9	5	7	21
Line 14	6	5	5	16
Line 15	7	4	7	18
Line 16	5	4	6	15
Total	116	67	78	

Fig. 32. Ranking of Lines (Poem B)

Verse Line	Total FG Features	Rank
Line 13	21	1
Line 11	19	2
Line 6	18	3
Line 15	18	4
Line 7	17	5
Line 12	17	6
Line 5	16	7
Line 9	16	8
Line 14	16	9
Line 16	15	10
Title	14	11
Line 2	14	12
Line 10	14	13
Line 1	13	14
Line 4	13	15
Line 3	11	16
Line 8	9	17

Fig. 33. Division of Lines (Poem B)

Verse Line	Rank	FG/BG			FG	BG
Line 13	1	FG		No. of lines	8	9
Line 11	2	FG				
Line 6	3	FG				
Line 15	4	FG				
Line 7	5	FG				
Line 12	6	FG				
Line 5	7	FG				
Line 9	8	FG				
Line 14	9	BG				
Line 16	10	BG				
Title	11	BG				
Line 2	12	BG				
Line 10	13	BG				
Line 1	14	BG				
Line 4	15	BG				
Line 3	16	BG				
Line 8	17	BG				

4.2.5. SATO'S *MIDAREGAMI* NO. 145

The tables below show all relevant FG totals and line rankings for this text.

Fig. 34. Table of FG (Poem C)

	Phonology	Syntax	Semantics	Total
Tone-unit 1	11	0	7	18
Tone-unit 2	2	2	4	8
Tone-unit 3	4	3	5	12
Tone-unit 4	8	4	5	17
Tone-unit 5	9	4	6	19
Total	34	13	27	

Fig. 35. Ranking of Tone-Units (Poem C)

Tone-unit	Total FG Features	Rank
Tone-unit 5	19	1
Tone-unit 1	18	2
Tone-unit 4	17	3
Tone-unit 3	12	4
Tone-unit 2	8	5

Fig. 36. Division of Tone-Units (Poem C)

Verse Line	Rank	FG/BG			FG	BG
Tone-unit 5	1	FG		No. of tone-units	2	3
Tone-unit 1	2	FG				
Tone-unit 4	3	BG				
Tone-unit 3	4	BG				
Tone-unit 2	5	BG				

4.2.6. BEICHMAN'S *MIDAREGAMI* NO. 145

The tables below show all relevant FG totals and line rankings for this text.

Fig. 37. Table of FG (Poem G)

	Phonology	Syntax	Semantics	Total
Line 1	7	1	8	16
Line 2	6	1	4	11
Line 3	7	0	3	10
Line 4	7	4	5	16
Line 5	8	3	4	15
Total	35	9	24	

Fig. 38. Ranking of Lines (Poem G)

Verse Line	Total FG Features	Rank
Line 1	16	1
Line 4	16	2
Line 5	15	3
Line 2	11	4
Line 3	10	5

Fig. 39. Division of Lines (Poem G)

Verse Line	Rank	FG/BG			FG	BG
Line 1	1	FG		No. of lines	2	3
Line 4	2	FG				
Line 5	3	BG				
Line 2	4	BG				
Line 3	5	BG				

5. EMPIRICAL RESULTS AND DISCUSSION OF FINDINGS

5.1. INTRODUCTION

In this section I discuss the results of the surveys. I have organised results by specific hypothesis tested (H_1 through H_{10}). I present a table summarising the experimental design for ease of reference below: it shows which surveys tested which hypotheses, which test instruments were employed, what kind of data these tests produced, and which statistical tests were used to analyse the data. At the end of this chapter, I also present a 'speed read' summary of conclusions drawn from the analyses, paired with the evidence—both quantitative and qualitative—that supports those conclusions. The conclusions in the 'speed read' table are not confined to H_1 through H_{10}; the table includes other insights which relate to the phenomena under discussion, but which fall outside the area of the enumerated hypotheses. These appear in the table under the category 'Other.' I draw upon these findings in the final chapter, where I discuss their broader significance.

Fig. 40. Experimental Design Summary

Hypothesis	Postulate	Survey	Instrument	Data	Stat. Test(s)
H_1	Strikingness	1 (1.1, 1.2)	Underlining test	***Nominal:*** two categories (words either underlined or not underlined) ***Ordinal:*** ranking of lines by no. of words underlined	Binomial, Chi-square test of independence, Spearman rank
H_2	Literary training	1 (1.1, 1.2)	Reader background questionnaire, Underlining test	***Nominal:*** two groups (familiar w/ theory vs. unfamiliar w/ theory) ***Ordinal:*** three groups (low, middle, high) ***Ordinal:*** amount of FG identified	Mann-Whitney *U*, Kruskal-Wallis, Friedman
H_3	Poetry Attitudes	1 (1.1, 1.2)	Likert scales	***Ordinal:*** summed scores from attitude questionnaire ***Ordinal:*** amount of FG identified	Kruskal-Wallis, Spearman rank
H_4	Age	1 (1.1, 1.2)	Reader background questionnaire, Underlining test	***Ordinal:*** age recorded in class intervals ***Ordinal:*** amount of FG identified	Kruskal-Wallis, Spearman rank
H_5	Impression of literariness	1 (1.1, 1.2)	Likert scales, Underlining test	***Ordinal:*** amount of FG identified ***Ordinal:*** summed scores from Likert scales	Spearman rank
H_6	Enjoyment	1 (1.1, 1.2)	Likert scales, Underlining test	***Ordinal:*** amount of FG identified ***Ordinal:*** summed scores from Likert scales	Spearman rank, Mann-Whitney *U*
H_7	Impression of foreignness	2 (2.1, 2.2)	Likert scales, Underlining test	***Ordinal:*** summed scores from Likert scales ***Ordinal:*** number of items underlined for impression of foreignness ***Nominal:*** two text conditions (ostensibly foreignising vs. ostensibly domesticating)	Rank biserial, Friedman, Spearman rank
H_8	Foreignising features	2 (2.1, 2.2)	Underlining test	Qualitative data	Qualitative analysis
H_9	Viability of foreignisation	2 (2.1, 2.2)	Likert scales	***Nominal:*** two text conditions ***Ordinal:*** summed scores from Likert scales	Mann-Whitney *U*
H_{10}	Translated-ness	3 (3.1, 3.2)	Likert scales	***Nominal:*** two text conditions ***Ordinal:*** summed scores from Likert scales	Mann-Whitney *U*

Note: the above statistical tests have been nonparametric, as opposed to parametric, in

nature. As Xin Gao (2010: 915) writes: "Nonparametric statistics refer to methods of measurement that do not rely on assumptions that the data are drawn from a specific distribution." The choice between parametric and nonparametric tests is an important one: "In practice, when the normality assumption on the measurements is not satisfied, parametric statistical methods might provide misleading results. In contrast, nonparametric methods make much less stringent distributional assumptions on the measurements" (ibid.). However, there is a cost associated with this benefit: nonparametric tests have less 'power' (Vaughan 2001: 153). That is, they are "less likely" than parametric tests to detect an effect "when in fact there is one" (ibid.). Nevertheless, when we take into account the underlying distributions, sample sizes and measurement levels of the data in my study, the choice is clear: nonparametric tests are more appropriate.

5.2: RESULTS

5.2.1: HYPOTHESIS 1 RESULTS (STRIKINGNESS)

H_1 concerns strikingness, a variable tested as a parameter of FG. The alternative hypothesis (H_a) predicted that words and phrases appearing in a foregrounded condition, as indicated by stylistic analysis, would be perceived as more striking. The null hypothesis (H_0) stated that whether text was foregrounded or not would make no significant difference in the reader's identification of strikingness. Respondents in Survey 1 knew that the texts were translations from the Japanese. I have used my stylistics analyses to put verse lines in order of 'containing the most FG' to 'containing the least FG.' Then I have used the frequency data from my surveys to produce another ranking, and finally I have compared the two sets of rankings to see how closely they agree or disagree.

Results for Survey 1[51] show:

- when we analyse all respondents together, **Poems A, C, E and F show significantly more underlined words in FG lines than in BG lines** (according to a binomial test)
  - *this is evidence **in favour of** the theory of foregrounding*
- however, when we analyse all respondents together, **results for Poems B and G show significantly more underlined words in *BG lines than in FG lines*** (according to a binomial test)
  - *this is evidence **against** the theory of foregrounding*
- when we break down results into Student and Non-Student groups, **we find their**

---

[51] Note: only poems A, B and C (from Survey 1.1) and poems E, F and G (from Survey 1.2) underwent stylistic analysis. This is because poem D forms a pair with poem H (my own translation) in Survey 1.2. In the interest of impartiality, I did not perform a stylistic analysis of these two poems (D and H): it seemed likely that my having written poem H could have affected my stylistic analysis of that text, thereby also affecting the comparison of Poems D and H along statistical lines.

patterns of response differ from each other for Poems B, C, E and G (according to a binomial test and chi-square test)

  o *this is evidence **against** the theory of foregrounding*

- finally, **predicted ranks of verse lines and observed ranks do not agree** for most poems (according to a Spearman rank correlation)

  o *this is evidence **against** the theory of foregrounding*

All in all, we have *not* found strong evidence in favour of the theoretical model for this hypothesis. In other words, there is little compelling evidence for us to reject the null model of reality—in the null model, the strikingness of a word or phrase has nothing to do with whether it appears in a foregrounded position in a poem. Moreover, the results partially confirm a suspicion I had in the initial stages of the project—that students and non-students would show different patterns of response. However, my original suspicion was that the *student group* would respond broadly as the theory predicted, and that only the *non-student group* would behave contrary to the expectations of the theory. This suspicion was founded on a notion discussed in Chapter 1: I supposed that the intersubjective agreement found in Van Peer's study (1986) might be due to the fact that all respondents were still university students (second- and third-year students at the University of Lancaster). It occurred to me that even though these students were divided into three groups along a "cline of expertise" (ibid.: 45), they might have already been trained (implicitly or explicitly) to respond in a basically similar way, regardless of which degree course they chose to pursue.

Taking a cue from De Certeau's (1984) essay "Reading as Poaching," I postulated that the longer a hypothetical reader had been out of university, the less of a 'hold' this institutional mode of reading might have over them, and therefore the greater the potential for diversity of response. De Certeau's essay is a defence of the autonomy of readers in the face

of "scriptural imperialism" (ibid.: 169)—a model of reading so labelled because it is exemplified for him by the pre-Reformation Church, which "ensured the Scriptures the status of a 'Letter' that was supposed to be independent of its readers" (ibid.: 172). As he puts it, "[i]n spite of the work that has uncovered an autonomy of the practice of reading underneath scriptural imperialism," the "*de facto* situation" in "contemporary culture" is the "assimilation of reading to passivity" (ibid.). In this 'scriptural' model of reading, to read "is to receive [a text] from someone else without putting one's own mark on it, without remaking it" (ibid.). Finally, what he terms 'scriptural' readings are those readings considered acceptable by the institutions in which they are produced. One such institution is the educational system. As De Certeau (1984: 171-172) writes:

> Reading is […] overprinted by a relationship of forces (between teachers and pupils, or between producers and consumers) whose instrument it becomes. The use made of the book by privileged readers constitutes it as a secret of which they are the "true" interpreters. […] Moreover, if the reader's expression of his freedom through the text is tolerated among intellectuals […], it is on the other hand denied students (who are scornfully driven or cleverly coaxed back to the meaning "accepted" by their teachers) […]. It is thus social hierarchization that conceals the reality of the practice of reading or makes it unrecognizable.

Thus, this survey was one way of testing the assertion that the "creativity of the reader grows as the institution that controlled it declines" (ibid.: 172). Since my experimental design did not ask respondents to compose their own interpretations, the approach of this survey might not seem like an ideal way of investigating the creativity (as such) of readers. However, if critical interpretation depends in part on which textual features stand out to readers, then we are justified in first investigating which features actually tend to grab readers' attention.

When we look at poem C, contrary to my initial suspicions, it is the *student group* that does not respond as predicted by the theory: there is no statistical relationship between underlining and the FG/BG categorisation of verse lines. Meanwhile, only the non-student group for this poem underlined enough words in the FG lines to lend support to the theory of FG. And when we look at Poem B, students' responses to the underlining task show a significant result in the *opposite direction to that predicted by H_a*. I mentioned above that for Poems B, C, E and G, student and non-student responses differ from each other. For Poem B, non-student responses show no significant relationship between line categorisation and underlining, while student responses do. For Poem C, non-student responses *do* show such a relationship, while student responses do not. For Poem E, students responded as the theory predicted, while non-students did not. And for Poem G, students underlined significantly more *BG items* than FG ones, while there was no relationship between line categorisation and underlining for non-students.

Readers may wish to refer to the tables below to see verse lines ranked by 'density' of FG, as well as the total number of items underlined by each respondent group per verse line. The term 'density' here refers to one of the "principal ingredients of the standard form of the theory of foregrounding" proposed by Van Peer (1986: 23-24). Density is the "vertical compaction of foregrounding devices on different linguistic levels in the texture of the poem at one particular point," and it is "achieved by the simultaneous occurrence of foregrounding devices on different linguistic levels, thus creating a nexus of foregrounding at one or more points in the text" (ibid.: 23). Van Peer contrasts 'density' with 'cohesion,' which he describes as a "horizontal force in the organization of the text: throughout the linear structure of the text a series of devices may be observed, displaying a number of similarities, thereby working in the same direction, and constraining the number of interpretative [sic] possibilities" (ibid.). While Van Peer's graphical representations of FG indicate both density

and cohesion, he writes that it is "obvious that text locations with a high density of FG should be given greater weight in the final quantification of FG" (ibid.: 68). Thus it was "mainly on the basis of such superimposed density that decisions concerning an item belonging to either FG or BG had to be taken" (ibid.). But it is not just density that influences the weighting in Van Peer's categorisation of lines as FG or BG. The level of description (phonology, syntax, semantics) also matters. Indeed, Van Peer assumes that FG on the phonological level will be "less conspicuous" than that on the grammatical level, which will in turn be "less prominent" than FG on the semantic level (ibid.: 68). However, the precise criteria and procedure for the weighting are left unexplained. Nevertheless, if we simply tally all the FG features for a particular poem in Van Peer's study—say, "Mirage" by Christina Rossetti—counting every symbol in the graphical representation (see Fig. 3.4 Visual Representation of FG in Rossetti's Poem in Van Peer [1986: 86]) as one instance of FG, then the resulting ranking corresponds to a high degree to the weighted ranking given by Van Peer (ibid.: 87). (A Spearman rank correlation identified a strong positive correlation between the two rankings, such that *rho* = 0.814 and p = .001.) Thus, in the absence of another method for assigning ranks to verse lines in my surveys, I have performed the same kind of tallying as above.

The tables below show verse lines for each stimulus text ranked by 'density' of FG, as well as the total number of items underlined by each respondent group per verse line. (Note: tables for Poems D and H are also included for reference. However, since I performed no stylistic analysis of these poems, three columns are missing: FG/BG, Total FG, and Mean Rank.)

Fig. 41. Line Rankings: Poem A

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
BG	Title	12	10.5	0	12	0	12	0	11
BG	Line 1	16	6.5	29	4	20	5	9	5
BG	Line 2	14	8	12	7	6	9	6	7.5
FG	Line 3	24	1	11	8	9	6.5	2	9
FG	Line 4	20	3	39	2	24	2.5	15	3
BG	Line 5	13	9	5	11	5	10	0	11
FG	Line 6	19	4	28	5	21	4	7	6
FG	Line 7	17	5	9	9.5	9	6.5	0	11
BG	Line 8	12	10.5	24	6	8	8	16	2
BG	Line 9	8	12	35	3	24	2.5	11	4
FG	Line 10	23	2	9	9.5	3	11	6	7.5
FG	Line 11	16	6.5	60	1	27	1	33	1
	Total	194		261		156		105	

Fig. 42. Line Rankings: Poem B

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
BG	Title	14	12	18	10	3	15	15	5
BG	Line 1	13	14.5	50	3	23	3	27	2
BG	Line 2	14	12	8	14	4	13.5	4	14
BG	Line 3	11	16	12	13	12	7.5	0	16.5
BG	Line 4	13	14.5	20	8	12	7.5	8	11
FG	Line 5	16	8	18	10	9	10	9	10
FG	Line 6	18	3.5	53	2	31	2	22	4
FG	Line 7	17	5.5	7	15.5	4	13.5	3	15
BG	Line 8	9	17	6	17	6	11.5	0	16.5
FG	Line 9	16	8	23	7	16	5	7	12
BG	Line 10	14	12	75	1	44	1	31	1
FG	Line 11	19	2	7	15.5	1	17	6	13
FG	Line 12	17	5.5	37	4	13	6	24	3
FG	Line 13	21	1	13	12	2	16	11	8.5
BG	Line 14	16	8	18	10	6	11.5	12	7
FG	Line 15	18	3.5	25	6	11	9	14	6
BG	Line 16	15	10	30	5	19	4	11	8.5
	Total	261		420		216		204	

Fig. 43. Line Rankings: Poem C

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
FG	Tone-unit 1	18	2	47	1	22	2	25	1
BG	Tone-unit 2	8	5	13	5	9	4	4	4.5
BG	Tone-unit 3	12	4	31	3	27	1	4	4.5
BG	Tone-unit 4	17	3	15	4	2	5	13	3
FG	Tone-unit 5	19	1	40	2	19	3	21	2
	Total	74		146		79		67	

Fig. 44. Line Rankings: Poem D

Verse Line #	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
Title	0	22	0	20.5	0	19.5
Line 1	11	8	3	14	8	5
Line 2	35	2	24	2	11	2
Line 3	2	20	2	16.5	0	19.5
Line 4	5	15	3	14	2	14
Line 5	13	7	8	5.5	5	11
Line 6	18	3.5	6	9	12	1
Line 7	15	6	9	4	6	8.5
Line 8	8	11	2	16.5	6	8.5
Line 9	18	3.5	8	5.5	10	3
Line 10	10	9	7	7.5	3	12.5
Line 11	16	5	10	3	6	8.5
Line 12	6	13	0	20.5	6	8.5
Line 13	1	21	1	18	0	19.5
Line 14	8	11	0	20.5	8	5
Line 15	3	18.5	0	20.5	3	12.5
Line 16	5	15	4	11.5	1	15.5
Line 17	4	17	4	11.5	0	19.5
Line 18	5	15	5	10	0	19.5
Line 19	3	18.5	3	14	0	19.5
Line 20	8	11	7	7.5	1	15.5
Line 21	37	1	29	1	8	5
Total	231		135		96	

Fig. 45. Line Rankings: Poem E

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
BG	Title	7	8	1	8	1	8	0	8
BG	Line 1	17	6	12	6	7	6.5	5	6
FG	Line 2	30	1	11	7	7	6.5	4	7
BG	Line 3	15	7	71	1.5	28	3	43	1
FG	Line 4	29	2	54	3	42	1	12	5
BG	Line 5	19	5	43	4	18	4	25	3
FG	Line 6	23	4	34	5	14	5	20	4
FG	Line 7	25	3	71	1.5	32	2	39	2
	Total	165		297		149		148	

Fig. 46. Line Rankings: Poem F

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
BG	Title	11	10	2	16	2	14	0	16.5
BG	Line 1	11	10	26	8	15	8	11	10
BG	Line 2	8	15.5	21	11	19	4	2	15
BG	Line 3	9	13.5	32	5	17	5	15	7.5
BG	Line 4	10	12	6	15	0	13	6	13
FG	Line 5	16	3	23	9.5	10	9	13	9
FG	Line 6	25	1	28	6.5	18	4	10	11
BG	Line 7	9	13.5	16	14	0	10.5	16	6
BG	Line 8	4	17	0	17	0	10	0	16.5
BG	Line 9	11	10	67	1	20	2.5	47	1
BG	Line 10	8	15.5	18	13	9	7	9	12
FG	Line 11	13	7.5	28	6.5	13	6	15	7.5
FG	Line 12	15	4	42	3	20	2	22	4
FG	Line 13	14	5.5	23	9.5	6	5	17	5
FG	Line 14	13	7.5	41	4	15	2.5	26	3
FG	Line 15	23	2	52	2	14	3	38	2
FG	Line 16	14	5.5	19	12	15	2	4	14
	Total	214		444		193		251	

Fig. 47. Line Rankings: Poem G

FG/BG (based on Total FG column)	Verse Line #	Total FG	Mean Rank (based on column at left)	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
FG	Line 1	16	1.5	49	2	19	2	30	1
BG	Line 2	11	4	51	1	25	1	26	3
BG	Line 3	10	5	15	5	6	4.5	9	5
FG	Line 4	16	1.5	30	4	6	4.5	24	4
BG	Line 5	15	3	43	3	16	3	27	2
	Total	68		188		72		116	

Fig. 48. Line Rankings: Poem H

Verse Line #	No. words underlined (all respondents)	Rank (based on all respondents)	No. words underlined (students)	Rank (based on student responses)	No. words underlined (non-students)	Rank (based on non-student responses)
Title	0	21	0	20	0	20
Line 1	27	5	11	6	16	4.5
Line 2	35	3	19	3	16	4.5
Line 3	9	11.5	7	9	2	16
Line 4	23	8	9	7.5	14	7
Line 5	9	11.5	6	11	3	14
Line 6	3	17	3	15	0	20
Line 7	29	4	14	4	15	6
Line 8	24	7	12	5	12	8.5
Line 9	8	14.5	4	14	4	13
Line 10	2	18	0	20	2	16
Line 11	8	14.5	0	20	8	10
Line 12	48	2	26	2	22	1.5
Line 13	8	14.5	6	11	2	16
Line 14	10	10	5	13	5	12
Line 15	18	9	6	11	12	8.5
Line 16	8	14.5	2	16	6	11
Line 17	0	21	0	20	0	20
Line 18	0	21	0	20	0	20
Line 19	1	19	1	17	0	20
Line 20	26	6	9	7.5	17	3
Line 21	62	1	40	1	22	1.5
Total	358		180		178	

Of course, it is difficult to see where statistical patterns lie just by looking at raw data. So I have included the additional tables below. They contain the relevant *p*-values for both the binomial and chi-square tests; although, as we shall see later, these results will come with a few caveats. Statistically significant results (*p*-values) that ***support the theory of foregrounding*** are highlighted in green. Results (*p*-values) that ***do not support the theory of foregrounding*** are highlighted in yellow. Results (*p*-values) that show a statistically significant relationship in the ***opposite direction of that predicted by the theory of foregrounding*** are highlighted in red. These red results not only ***do not support the theory of***

foregrounding, but also suggest that the relationship between FG and strikingness in the theory of foregrounding is backwards, i.e that BG lines are more likely to have items underlined for strikingness.

Fig. 49. Strikingness Results (Poem A)

Categ ory	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value
FG	156	0.002	2.596	0.107	93	0.020	1.479	0.224	63	0.050	1.125	0.289
BG	105				63				42			
	261				156				105			

Fig. 50. Strikingnness Results (Poem B)

Categ ory	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value
FG	183	0.010	0.572	0.449	87	0.005	2.501	0.114	96	0.441	0.323	0.570
BG	237				129				108			
	420				216				204			

Fig. 51. Strikingness Results (Poem C)

Categ ory	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ^2	χ^2 test P-value
FG	87	0.025	0.231	0.631	41	0.822	1.542	0.214	46	0.003	4.368	0.037
BG	59				38				21			
	146				79				67			

Fig. 52. Strikingness Results (Poem E)

Category	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value
FG	170	0.015	8.824	0.003	95	0.001	0.058	0.809	75	0.935	15.288	0.000
BG	127				54				73			
	297											

Fig. 53. Strikingness Results (Poem F)

Category	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value
FG	256	0.001	29.331	0.000	111	0.044	7.659	0.006	145	0.016	10.516	0.001
BG	188				82				106			
	444				193							

Fig. 54. Strikingness Results (Poem G)

Category	All Respondents				Student Respondents				Non-Student Respondents			
	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value	Words under-lined	Binomial Test P-value (2-tail.)	χ²	χ² test P-value
FG	79	0.034	16.420	0.000	25	0.013	15.251	0.000	54	0.516	4.228	0.040
BG	109				47				62			

There are two reasons to interpret the above results with caution. The first concerns the binomial test. The binomial test determines whether the difference in frequency of words underlined in the FG and BG lines is statistically significant, based on an expected proportion of 0.50 in each category. This reflects the idea that if the recorded frequencies were the result of chance (and not the result of the independent variable, a.k.a. foregrounding), then the

number of words underlined in each text condition—i.e. in an FG or BG line—should be "approximately half of the total number of observations" (Van Peer 1980: 540). But in this case, the application of this test rests upon an idiosyncratic understanding of the notion of the experimental 'trial.' Ogunnaike (2010: 59) defines a trial as "a single performance of a well-defined experiment giving rise to an outcome." The classic example is a coin toss. A coin toss is a single performance; it normally yields one of two outcomes (heads or tails); and if we want to test whether the proportion of each outcome is 0.50, then we have a falsifiable hypothesis and thus the experiment is well-defined. It seems possible to consider the *whole activity* of one respondent underlining items in a poem as a single trial, since it results in a single outcome (total number of items underlined). But the outcomes that are actually being observed here are outcomes *per line* (or tone-unit, in the case of Sato's translation).[52] The problem is: the binomial test assumes that each trial has the same probability of success; in other words, no trial has an influence on any successive trial. But would this really be the case as a respondent moves from one line to the next in a poem? Especially considering that not all lines are equal length, and that they terminate at different points in a sentence? In other words, would the probability of each outcome in the null model really 'reset' itself to 0.50?

The second thing to note concerns the chi-square test. Unlike the binomial test, which deals with frequencies of underlined words in FG lines, the chi-square test also accounts for *words that were not underlined*. In other words, it examines the "relationship between […] words being underlined or not underlined in the responses and their belonging to either the FG or BG category" (Van Peer 1986: 100). Because it uses more data, it should be a more stringent test, and therefore should make our predictions more "vulnerable to falsification"

---

[52] Since Sato's translation was printed in my survey as one continuous line, and statistical analysis of FG required the text to be broken into units, I chose to analyse the text according to the 5 tone-units identified in my stylistic analysis. The rationale: the tone-unit boundaries corresponded to the poem's punctuation marks, breaking the poem up into five visible units. (In the end, underlinings from the strikingness test rarely spanned more than one tone-unit, suggesting that Sato's text was indeed perceived along these lines.)

(Van Peer 1980: 541). However, the application of this test rests on the assumption that lines belonging to the FG category would have proportionally more items underlined in them than BG lines. Yet a long line of verse could have a relatively large number of FG features which nevertheless involve only a few items in that line—the FG could be very dense around, say, only 2 out of 10 words in a particular line. Density of FG would result in that line's promotion to a high rank in terms of FG. However, if respondents behaved as H_a predicts, then that same line would also have a relatively high number of *not* underlined words. Indeed, this seems to have happened in the non-student results for Poem E: while the binomial test reported no association between verse line category (FG/BG) and number of words underlined, the chi-square test *does report* a strong association between the two variables. While the number of words underlined in FG lines is close to that for BG lines (75 and 73, respectively), the number of words *not underlined* for both types of verse lines is quite different. More than twice as many words in FG lines were not underlined than in BG lines. Thus FG lines were more likely to contain *not underlined words* than BG lines. Because this is broadly the opposite of what the theory predicts, yet the result is significant, I have shaded the appropriate cell red. A similar situation holds for Poem G: there were significantly more words *not underlined* in FG lines than in BG ones. This argues for a more qualitative approach which accounts for the relative contribution of *each underlined item* to the overall strikingness of the line.

It is also worth comparing respondents' rankings of the verse lines (based on total words underlined) with my own rankings of the lines (based on total FG features identified per line). This can be done with a Spearman rank-order correlation. The values to look out for are not just the *p*-values but also the correlation coefficients, which range from -1.00 to 1.00. (Due to limitations of space, I present only the table for Poem F below).

For only one of the poems (Poem F) is there any significant agreement between my

rankings and the respondents' rankings, whether they are analysed in subgroups or all together as one big group. The results for Poem F show a moderate, significant agreement between my rankings and those of respondents when seen as one big group. The next closest thing to a match with my rankings occurs with Poem C: there is a high correlation between my rankings and non-student respondents (0.872), as well as between my rankings and all respondents (0.800), but the *p*-values do not fall in the appropriate region (they are 0.054 and 0.104, respectively). As the insignificant results for the other poems outweigh these cases, this is evidence *against the theory of FG*. Meanwhile, for poems A, B, E and H, student and non-student rankings show a moderate-to-high degree of correspondence with each other at a significant level (their correlation coefficients are 0.697, 0.502, 0.731, and 0.823, respectively; the associated *p*-values are: 0.012, 0.040, 0.040, and 0.000). This *would* be evidence in favour of the idea that strikingness depends less on the reader than on the text. However, for the other four poems, there is no significant agreement between rankings from students and rankings from non-students.

Fig. 55. Spearman Rank Correlations: Poem F

			Experimenter	All Respondents	Students	Non-Students
Spearman's rho	Experimenter	Correlation Coefficient	1.000	.539*	.290	.435
		Sig. (2-tailed)	.	.026	.260	.081
		N	17	17	17	17
	All Respondents	Correlation Coefficient	.539*	1.000	.750**	.824**
		Sig. (2-tailed)	.026	.	.001	.000
		N	17	17	17	17
	Students	Correlation Coefficient	.290	.750**	1.000	.397
		Sig. (2-tailed)	.260	.001	.	.115
		N	17	17	17	17
	Non-Students	Correlation Coefficient	.435	.824**	.397	1.000
		Sig. (2-tailed)	.081	.000	.115	.
		N	17	17	17	17

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

As the results stand now, we cannot really say that the theory of FG has allowed us to predict

which verse lines would attract the most attention. This is evidence *against the theory of FG.*

H_2 concerns the relationship between literary training and the identification of strikingness. H_a predicted that the more literary training readers indicated, the more FG they would identify (in terms of number of words underlined for strikingness). H_0 stated that the amount of literary training would *not* be significantly correlated with the amount of FG underlined. The test instruments used to collect data were the underlining test (from H_1), as well as two reader background questions from the questionnaire: 1) 'Are you familiar with the theory of foregrounding?' and 2) 'Have you ever studied literature or poetry in a class/workshop/reading group? If so, where?' The statistical tests employed here are the Mann-Whitney U test and the Spearman rank test. The Mann-Whitney U test is "often used to compare two groups with regard to a given criterion that is measured on at least an ordinal scale" (Pett 1997: 169). In this case, the two groups are respondents who answered 'yes' to the first question above, and those who answered 'no.' The Kruskal-Wallis test is "an extension of the Mann Whitney test that can be applied to three or more" groups (McKillup 2012: 331). Here I divided respondents into three groups based on their responses to the second question above: High, Mid and Low. (Respondents who had at least university-level experience with literature or poetry were assigned to the High category; those with at least high school-level experience were assigned to the Mid category; and those who had not studied literature or poetry at either university- or high school-level were assigned to the Low category.) Finally, I perform a Friedman two-way ANOVA to determine whether there is any relationship between experience level and verse lines with regard to amount of FG indicated. The Friedman test can help determine whether there are *qualitative* (rather than just quantitative) differences in the perception of strikingness among respondents. A significant result in either the Mann-Whitney or Kruskal-Wallis test should be interpreted as evidence

against the theory of foregrounding. That is, the theory *prefers the null hypothesis (H₀)*, and a significant result on either of those tests would mean rejecting H$_0$. Meanwhile, a significant result on the Friedman test should be taken as evidence *in favour of* the theory, insofar as this would suggest that respondents largely agreed in their assessment of the verse lines, regardless of their experience level.

Results for Survey 1 show:

- **there is no significant relationship between amount of FG identified per poem and familiarity with the theory of FG**, for any of the poems (according to the Mann-Whitney U test)
    - *this is evidence **in favour of** the theory of FG*

- **there *is* a significant correlation between training level and amount of FG identified for Poem D** (according to the Kruskal-Wallis test)
    - *this is (weak) evidence **against** the theory of FG*

- **there is no significant correlation between training level and amount of FG identified per poem for Poems, A, B, C, E, F, G and H** (according to the Kruskal-Wallis test)
    - *this is evidence **in favour of** the theory of FG*

- **there *is* a significant relationship between training level and amount of FG identified *per line* for Poems B, C and G** (according to the Friedman test)
    - *this is evidence **against** the theory of FG*

Overall, we have found some good evidence in favour of the theory of FG, but we have also found some evidence against it. First the results in favour of the theory of FG: it appears that for these respondents, for this set of stimulus texts, there is no relationship

275

between familiarity with the theory and the amount of words underlined per poem for strikingness. *This broadly supports the notion that it is the text—rather than the expertise of the reader—that determines response.* For 7 out of the 8 texts, there also appears to be no relationship between training level and amount of FG underlined per poem. Again, this result supports the theory of FG.

However, it is possible that the number of items underlined per poem was influenced by the instructions of the questionnaire: these said to underline 'up to a maximum of ten words or phrases.' Thus, it might be the case that people will more or less identify the same amount of FG per poem regardless of their expertise level, at the same time that the *locations* of this underlined FG differ systematically. This qualitative issue was investigated via the Friedman test, which revealed that when respondents were grouped by their general level of experience with literature or poetry, they *sometimes did disagree* about which locations were the most striking. *As this result supports the conventionalist position, this is evidence against the theory of FG.*

As I summarised earlier, the Mann-Whitney U tests revealed no significant relationship between amount of FG underlined and familiarity with the theory of FG for any of the poems. The relevant statistics and *p*-values can be seen in the table below. The null hypothesis for the Mann-Whitney test is that the two groups (familiar with theory vs. not familiar with theory) come from the same population. This would imply that familiarity with the theory makes no significant difference on the underlining test. As the theory of FG prefers the null hypothesis, *p*-values above 0.05 are evidence in favour of the theory. I have highlighted *p*-values that *support* the theory of FG in green. The columns for mean and median reflect the mean and median number of words underlined per poem. N is the number of respondents.

Fig. 56. Mann-Whitney *U* Results—Strikingness vs. Familiarity with Theory

Poem	Not Familiar w/ Theory of Foregrounding			Familiar w/ Theory of Foregrounding			Mann-Whitney	
	N	Mean	Median	N	Mean	Median	U	p (exact 2-tailed)
Poem A: Strikingness (No. Words Underlined)	27	8.85	8.00	4	5.50	6.50	36.00	0.32
Poem B: Strikingness (No. Words Underlined)	26	15.04	12.50	3	9.67	13.00	31.50	0.61
Poem C: Strikingness (No. Words Underlined)	26	4.88	4.00	3	6.33	7.00	24.00	0.32
Poem D: Strikingness (No. Words Underlined)	25	8.28	9.00	4	6.00	5.50	36.00	0.41
Poem E: Strikingness (No. Words Underlined)	24	9.83	9.50	4	11.75	12.50	36.50	0.47
Poem F: Strikingness (No. Words Underlined)	24	15.88	14.50	4	13.25	13.00	45.50	0.87
Poem G: Strikingness (No. Words Underlined)	25	6.20	6.00	4	5.25	5.50	42.50	0.65
Poem H: Strikingness (No. Words Underlined)	24	13.04	12.00	4	10.00	7.00	39.50	0.59

Next, I discuss the results of the Kruskal-Wallis test for strikingness vs. training level. These results should be approached with caution, as the test requires 5 members in each category to work optimally, and this minimum number was not met for the Low group in Survey 1.1. Meanwhile, the Mid group did not reach 5 in Survey 1.2. The null hypothesis of the Kruskal-Wallis test is that the three groups come from the same population. This would imply that training level makes no significant difference on the results of the underlining test. As the theory of FG favours the null hypothesis, *p*-values above 0.05 are evidence in favour of the theory. In fact, the *p*-values for all the poems except Poem D were above 0.05.

Since the model presented in Van Peer (1986) holds that experience levels should *not* affect responses, the results from this test broadly *support* the theory of FG: apart from the results for Poem D, *there are no significant differences between the three groups of respondents with regard to amount of FG underlined. That is, training level and amount of FG identified are not associated.* It is difficult at this time to know why Poem D—and only Poem D—shows a significant result for the Kruskal-Wallis test. But it is useful to remember that this test compares mean ranks, and the group with the highest mean rank contains the most top-scoring respondents. By 'top-scoring' I mean those respondents who underlined the most words. In the case of Poem D, that group is the Mid group. This does *not* conform to the

prediction that the more literary training respondents have, the more FG they will underline. Nevertheless, as the results for this poem show something *other than* the hypothesis preferred by the theory (that there will be no relationship between literary training and strikingness), it is evidence *against* the theory, albeit relatively weak evidence.

So: is there something special about Miller and Kudo's text ("Instant") that would make its language stand out more frequently to respondents in the Mid group, as opposed to the Low and High groups? With its frequent repetitions of lexis and syntax, this text arguably contains the most parallelism of any in the corpus. Could it be that respondents in the Mid group tend to find parallelism more striking than respondents in the other two groups? If so, then we might expect the Mid group to have the highest mean rank when it came to other texts with a lot of repetition. My translation (Poem H), for instance, also frequently repeats lexis and syntax. However, the Mid group has the *lowest* mean rank for that text. Bownas and Thwaite's text (Poem B) also displays a lot of repetition of lexis and syntax, yet it is the Low group that scores the highest mean rank for that text. This variability in response among the groups raises an interesting question: if the foregoing results broadly suggest that there is no significant association between training level and amount of FG underlined *per poem*, can it also be demonstrated that there is no association between training level and amount of FG underlined *per line*? This is an important question to ask insofar as it shades into a qualitative (rather than quantitative) issue.

While it is possible that people will more or less identify the same amount of FG per poem regardless of their expertise level, it is also possible that the *locations* of this FG may differ systematically. It would not be unreasonable to suspect that such differences might themselves reveal patterns that we could further investigate: do readers with the most training tend to notice a particular type of FG feature over others? And so on. If such evidence were found, it would go a long ways in corroborating the conventionalist intuition that what we

notice in a text—the features on which we base our interpretations—are influenced by the particular interpretive communities to which we belong (see Fish 1980b).

Moreover, when we consider that these texts are translated, and that understandings of translation may differ *within* interpretive communities, such evidence might also complexify the very notion of interpretive communities. As Scholes (1984: 174) has pointed out, in Fish's theory "no one can belong to more than one" interpretive community at a time. The relationship between these communities is always disjunct: "members of different communities will disagree because from each of their respective positions the other 'simply' cannot see what is obviously and inescapably there" (Fish 1980a: 15). This is because Fish assumes that *all* members of a particular community will 'see' (and therefore construct) the same text in the *same* way, and that this text will have nothing in common with the same text as 'seen' (constructed) by another community—at least in terms of what *really* counts (ibid.). In Fish's theory, readers can move *between* interpretive communities; however, as Scholes (1984: 174) writes, "members of the same interpretive community, by definition, have no disagreement." In insisting too much on the "stability" of interpretive communities (Fish 1980a: 15) as the principle that allows different readers to perceive different "formal features" (ibid.: 14), Fish not only overlooks the *relative* stability of the written text at any given historical-linguistic juncture (cf. Easthope 1991: 33-42), but he forecloses the possibility of diversity of response *within* single communities. Fish ([1976] 1980: 171-172) acknowledges that all interpretive communities are "temporary," but in his model, as long as they exist, they have a kind of fixed identity. Thus, coming to an agreement with someone involves leaving one group for another. What Fish fails to see is how individual readers might introduce *changes* to their respective interpretive communities; he does not see how agreement may be explained by realignments or partial *overlaps between* interpretive communities. If the translated-ness of texts triggers different reading behaviours among

readers who otherwise agree, then translation indeed "open[s] up the plural signifying potential" not just of the source text, as Reynolds (2019: 3) argues in *Prismatic Translation*, but of the TT as well. Thus we might take a 'prismatic' view of interpretive communities; this would involve seeing them in much the same way that Reynolds (ibid.) considers languages: "more as a continuum of variation than as a collection of bounded entities."

A Friedman two-way ANOVA can help us investigate qualitative differences in response. While our Kruskal-Wallis test analysed how the amount of underlined FG (the dependent variable) differed as a function of experience level (the independent variable), the Friedman test can analyse how the amount of underlined FG differs as a function of both experience level and individual verse line. In other words, it can help us determine whether certain verse lines consistently attract more attention than others *across the three expertise levels.*

The results of the Friedman tests for each poem are presented in the table below. The null hypothesis of the Friedman test is that verse lines will have about equal mean ranks across the experience groups. While the phrase 'equal mean ranks' sounds like it should indicate agreement between groups, *in this case it actually implies the opposite*. That is, in the null model of reality, there is little consistency between the three subgroups concerning which verse lines were the most striking (in terms of number of words underlined), as every verse line would have about an equal chance of assuming the top spot among the three groups. In other words, in the null model, there is *disagreement* between the three groups about which verse lines are the most striking. The theory prefers the alternative hypothesis, that respondent groups *agree* about which verse lines are the most striking, since this would imply it is the text itself that determines the response. Thus, a significant *p*-value (under 0.05) *supports* the theory of FG, since it rejects the null and supports the alternative hypothesis. I have highlighted results that *support the theory of FG* in green results that *do not support the*

theory in <mark>yellow</mark>.

Fig. 57. Friedman ANOVA Results—Experience Level vs. Strikingness

	N (no. of subgroups)	Chi-square	df	p-value	Kendall's W
Poem A	3	22.84	11	0.019	0.69
Poem B	3	21.54	16	0.159	0.45
Poem C	3	8.276	4	0.082	0.69
Poem D	3	34.35	21	0.033	0.55
Poem E	3	17.56	7	0.014	0.84
Poem F	3	31.352	16	0.012	0.65
Poem G	3	8.339	4	0.080	0.70
Poem H	3	49.408	21	0.000	0.78

As we can see from the *p*-values for Poems B, C and G, the null hypothesis of equal mean ranks for the verse lines has been *retained* (since the *p*-values are above 0.05). This implies disagreement among the subgroups—that the groups differ in terms of which verse lines attracted the most underlining. *Differences in the underlining test among the three subgroups are not attributable solely to the verse lines themselves; rather, the experience level of the respondent is also implicated.* However, as the Friedman test is an omnibus test, it tells us only *whether* there are significant differences, but not where the differences lie (i.e., between which verse lines across the three subgroups). For that information, lengthy post hoc testing would be required. As an alternative method, Van Peer (1986: 118-120) presents line graphs featuring a separate line for the response of each subgroup. As Van Peer (ibid: 118) writes, qualitative differences between the groups "can be derived from the direction the graph takes after each point." If the polygons for two different groups go off in different

directions after a single coordinate, this would be construed as a qualitative difference in response. As for the valleys in the graphs, the deeper the valleys, the less evenly distributed the amount of underlined FG throughout the poem. As Van Peer (ibid.) writes, deep valleys for a particular subgroup can "be explained as these informants' tendency to make sharper discriminations between what is striking and what is not."

The three graphs below show the average number of words underlined per line, per group for Poems B, C and G. The X-axis displays both line number and line rank in terms of amount of FG identified in stylistic analysis. The Y-axis displays average number of words underlined. The data for Poems B and C come from the same respondents, as both texts appeared together in Survey 1.1. The data for Poem G, however, come from a different group of respondents (albeit largely sampled from the same places as readers of B and C). This might explain the similarities in the responses to Poems B and C: for both texts, the group that makes the sharpest distinctions in terms of strikingness is the Low group; also, except for a couple spots where they cross each other in Poem B, the trend lines for the Mid and High groups largely behave like each other, rising when the other rises, falling when the other falls. This suggests a qualitative difference in response between the Low group on the one hand, and the Mid and High groups on the other. It is tempting to argue (as I did in H_1) that by the time they reach university, respondents have already been trained (implicitly or explicitly) to respond in a basically similar way. In other words, there might be a baseline level of formal training, which once attained by respondents ensures a certain degree of qualitative similarity in terms of response; this level might be the high school level. After all, the criterion that set the Low and Mid categories apart was experience studying literature or poetry at the high school level. *This might be the cut-off point that explains why the Low group responded in a qualitatively different way.* As I wrote above, it is *tempting* to offer this explanation. But the results for Poem G complicate the picture. There, it is the trend lines for the *Low and High*

categories that start off similarly; meanwhile, the Mid and High trend lines disagree with each other for most of the graph, with one line rising as the other falls, and vice versa. Thus, we cannot straightforwardly accept the above explanation. Of course, whether we have another theory ready that can explain all three graphs does not change the situation: these graphs suggest qualitative differences in response attributable to training level. Therefore, they present evidence *against* the theory of FG.

Fig. 58. Poem B: Average No. Words Underlined
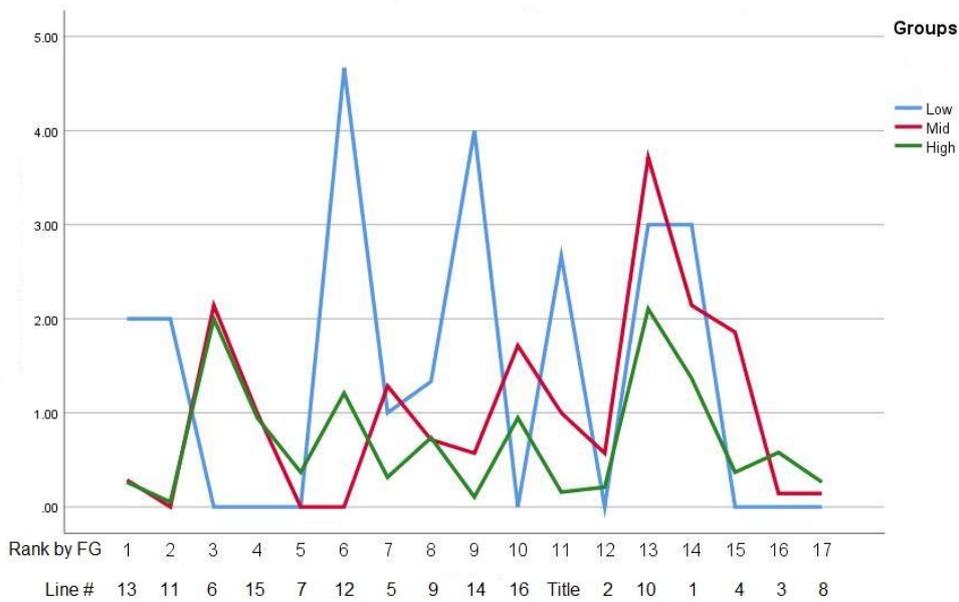
Fig. 59. Poem C: Average No. Words Underlined
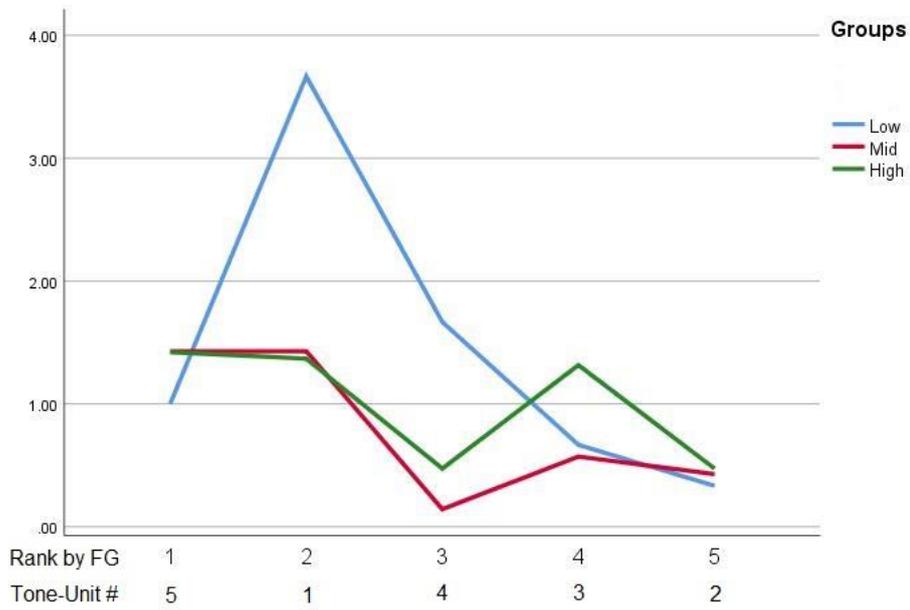
Fig. 60. Poem G: Average No. Words Underlined
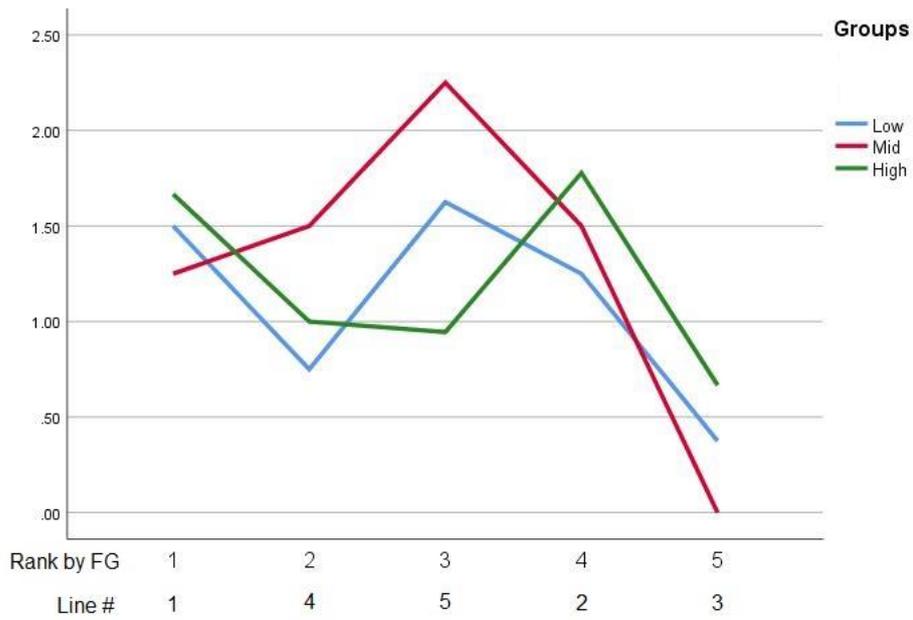
H_3 concerns whether attitudes toward poetry are related to the identification of strikingness. H_a predicted that the higher respondents scored on an attitude test, the more FG they would underline. H_0 stated that attitudes toward poetry would not be significantly correlated with amount of FG underlined. The test instruments used to collect data were the underlining test for strikingness (from H_1), as well as the 11 Likert-scaled items concerning poetry attitudes from the Reader Background page of the survey. The statistical tests employed here are the Kruskal-Wallis test and the Spearman rank correlation. These tests represent two different ways of approaching the question. The K-W test allows us to posit three groups based on responses to the attitude test: 'Low,' 'Middle' and 'High.' Then we can determine how likely it is that differences between the groups are down to chance.[53] The Spearman rank correlation, on the other hand, does not require us to break respondents into groups: it treats both poetry attitudes and amount of underlined FG as ordinal-level variables to determine whether any correlation exists between them. A significant result on either test should be interpreted as evidence *against* the theory of FG, since such a result would suggest that differences in poetry attitudes are related to differences in response to FG.

The results for Survey 1.1 and 1.2 show:

- **there is no association between poetry attitudes and the identification of FG (in terms of number of items underlined per poem)** for any of the texts (according to the Kruskal-Wallis test)
  - *this is evidence **in favour of** the theory of FG*
- **there is no correlation between poetry attitudes and the**

---

[53] Van Peer (1986: 122) uses a chi-square test here. However, one of the variables (poetry attitudes) has ordered levels. As the chi-square test is "insensitive to order" (Pett 1997: 197), I chose a test that is sensitive to order.

identification of FG (in terms of number of items underlined per poem) for any of the texts (according to the Spearman rank correlation)

      o   *this is evidence **in favour of** the theory of FG*

Results for this hypothesis supported the theory of FG. For the Kruskal-Wallis test, it was necessary to split respondents into groups based on their scores on the poetry attitude scale at the back of the surveys. I used the following ranges: Low (11 to 44), Mid (45 to 60) and High (61 to 77). The above ranges are not equal; indeed, defining the ranges for 'Low,' 'Mid' and 'High' is a task in which subjectivity comes into play. I was guided by two main considerations. One was having the minimal sample size per group for the Kruskal-Wallis test to work optimally; this value is 5 (Privitera 2012: 600). The second consideration had to do with the nature of the samples. Given that most of the data in this survey came from either students in literature classes or members of local poetry writing/appreciation groups, one would not expect to find many low scores on the whole. Rather, the mean and median scores for these respondents would probably be higher than for the population at large. In other words, because these respondents have presumably *chosen* to study literature of their own accord, either in a formal or informal setting, one would expect them to hold relatively positive attitudes toward poetry, with the poetry group members probably scoring the highest, as these respondents have chosen to engage with poetry beyond their formal education. Of course, it is possible for literature students to dislike poetry. However, on this point, we may turn to the results of the "Literature in Britain Today" survey (The Royal Society of Literature 2017a). This survey found that poetry was the second most common type of "reading material considered to be literature," coming in one percentage point behind novels (ibid.: 12). Moreover, the survey found that "88% of people agree that literature should be part of everyone's education" (ibid.: 3). If literature is thus "valued highly in British society" (ibid.)—and if poetry is relatively central to understandings of that category—then how likely

is it that literature students on the whole would have a lower opinion of poetry than the population at large? This supports the idea that there should be relatively few respondents assigned to the Low category. As we will see below, most respondents were assigned to either the Mid or High categories.

To ensure that at least 5 respondents fell into the 'Low' category, I had to set the boundary between 'Low' and 'Mid' at 44. This was the lowest I could set that boundary for the Kruskal-Wallis test to work well. As for the next boundary—the boundary between 'Mid' and 'High'—I chose 60. This seemed to accord well with the data: there seemed to be a 'jump' in scores around this point: while many scores fell between, say, 50 and 60, there were not many scores in the low 60s; scores seemed to shoot up from there to the mid and high 60s. Also, this was the number above which there appeared to be more poetry group members than students in the data. (We would probably expect poetry group members to be more likely to fall into a category labelled 'High' than students, as they have been motivated to continue to engage with poetry beyond [sometimes well beyond] the end of their formal education.)

The Kruskal-Wallis results appear in the table below. Once again, I have highlighted results that *support the theory of FG* in green. A significant result on this test (i.e. a *p*-value under 0.05) would indicate that the null hypothesis of equal mean ranks should be rejected; this would suggest that poetry attitudes *are* associated in some way with amount of underlined FG. Since the 'standard' model of the theory holds that poetry attitudes should *not* influence responses, a significant result on this test *would not support the theory of FG*. However, as we see below, all *p*-values are above 0.05, and are therefore highlighted in green. These results suggest that poetry attitudes and the identification of strikingness are not associated.

Fig. 61. Kruskal-Wallis Results—Poetry Attitudes vs. Strikingness

	Low N	Mid N	High N	Mean Rank (Low)	Mean Rank (Mid)	Mean Rank (High)	K-W H	df	p-value
Poem A	10	15	7	13.65	16.10	21.43	2.902	2	0.234
Poem B	10	13	6	14.65	12.85	20.25	3.139	2	0.208
Poem C	10	14	6	13.40	15.68	18.58	1.335	2	0.513
Poem D	8	15	7	13.88	16.63	14.93	0.555	2	0.757
Poem E	6	16	8	13.67	15.94	16.00	0.330	2	0.848
Poem F	6	16	8	15.00	15.16	16.56	0.162	2	0.922
Poem G	6	17	8	11.17	17.32	16.81	2.146	2	0.342
Poem H	5	15	7	12.40	14.90	13.21	0.468	2	0.791

Next I used the Spearman test to compare respondents ranked by attitude score with the same respondents ranked by number of items underlined for strikingness per poem. If either a positive or negative correlation existed between these two variables, it would show up as a significant result. Such a result would *not support the theory of FG* since in the 'standard' model, poetry attitudes and response should not be related. **However, no p-values were below 0.05.** *This suggests there was no correlation between poetry attitudes and the identification of strikingness.* In other words, holding a positive or negative attitude toward poetry made no significant difference in terms of the amount of FG identified per poem. This is evidence *in favour of the theory of FG.* (I omit tables here out of concerns for space.)

H_4 concerns the relationship between age and identification of FG (in terms of words underlined for strikingness). H_a stated that there will be a significant correlation between respondent age and number of words underlined. H_0 held that there will be no relationship between age and identification of strikingness. The test instruments used to collect data were the underlining test for strikingness (from H_1), as well as the question concerning age from the Reader Background page of the survey: there were six possible age groups to choose from (18 – 24, 25 – 34, 35 – 44, 45 – 54, 55 – 64 and 65+). As we saw above, the Kruskal-Wallis test works best when there are at least 5 respondents per group, and it is sometimes necessary to combine groups so that that number is met (so long as it is theoretically justifiable). In this case, I combined the 18-to-24 and 25-to-34-year-olds into one group, the 35-to-44 and 45-to-54-year-olds into another group, and the 55-to-64 and over-65-year-olds into another group, for a total of three groups: Low, Mid and High. The number of 5 respondents was still not met for the Mid group for any of the poems; however, this was preferable to having multiple categories of just 1 or 2 respondents. A significant result on this test would suggest that at least one group is different from the others in terms of amount of FG identified. This in turn would suggest that age and the identification of strikingness are associated in some way, which goes against the 'standard' model of the theory of FG.

Results for Surveys 1.1 and 1.2 show that:

- **there is a significant association between age group and amount of FG (no. of words) identified in Poem F, with the Mid group underlining the most** (according to the Kruskal-Wallis test)
  - *this is (weak) evidence **against** the theory of FG*
- **there are no significant associations between age group and amount of FG (no. of**

words) identified for any of the other texts (according to the Kruskal-Wallis test)

- o *this is evidence **in favour of** the theory of FG*

- **belonging to one of the above age groups does not make one more (or less) likely to underline striking words or phrases** (according to the Kruskal-Wallis test)

  - o *this is evidence **in favour of** the theory of FG*

- **nevertheless, for some poems, older respondents tended to make fewer underlinings** (according to the Spearman rank test)

  - o *this is evidence **against** the theory of FG*

- **and for some poems, the older the respondent, the more likely they were to perceive longer units of text as striking** (according to the Spearman rank test)

  - o *this is evidence **against** the theory of FG*

Results for this hypothesis are somewhat mixed, with the Spearman test results suggesting a possible difference between survey groups in the *manner in which strikingness is perceived*. First, I discuss the Kruskal-Wallis test results. I performed two kinds of K-W tests: one to check for associations between age and total number of words underlined per poem, and another the check for associations between age and number of individual underlinings (irrespective of how many individual words were underlined).[54] A significant result ($p < 0.05$) on this test would suggest that at least one of the groups is different from the others in terms of amount of FG identified. Such a result would *not support the theory of FG*, as the 'standard' model holds that age and response should not be associated. **There was only one significant result**—this is the result for Poem F (Age vs. Number of Words Underlined), shown in the table below. Since it does not support the theory

---

[54] A Spearman rank test revealed no significant correlations between either literary training or poetry attitudes and number of underlinings (as opposed to number of words underlined) in H_2 and H_3.

of FG, it is highlighted in <mark>yellow</mark>.

Fig. 62. Poem F: Kruskal-Wallis (Age vs. No. Words Underlined)

Ranks

	Age Group (Low, Mid, High)	N	Mean Rank
Number of Words Underlined	Low	15	14.27
	Mid	4	26.75
	High	11	13.09
	Total	30	

Test Statistics[a,b]

	Number of Words Underlined
Kruskal-Wallis H	7.719
df	2
Asymp. Sig.	.021

a. Kruskal Wallis Test

b. Grouping Variable: Age Group (Low, Mid, High)

It is difficult at this stage to say why Poem F—and only Poem F—should yield a significant result. As we can see from the graph below, Line 7 attracted considerable attention from the Mid group, while boasting no items underlined by either of the other groups. Line 7 runs: "But sometimes they like to have friends on Earth." Line 9 also attracted much more attention from the Mid group than from the others. Lines 9 and 10 run: "Universal gravitation is the power of solitudes / pulling each other." An inspection of these lines reveals that they both have relatively few FG features identified by my stylistic analysis—they were both classified as BG lines, with Line 9 ranked 11th and Line 7 ranked 14th out of 17 verse lines (including title) in terms of amount of FG. Furthermore, of the three different levels of analysis, both lines have more FG on the phonological level than on the other levels. Finally,

the assonance between "Earth" and "Universal" in the two lines constituted a parallelism. Could it be that the Mid group favours parallelisms on the phonological level? If this were the case, then we would expect to see a relatively high number of items underlined in other lines with such features, e.g. Lines 5-6 and 15-16. And yet the results do not really bear this out. Nevertheless, since age group *is* associated with *amount of FG* identified here, it constitutes (weak) evidence against the theory of FG.

Fig. 63. Poem F: Average No. Items Underlined Per Line, Per Age Group
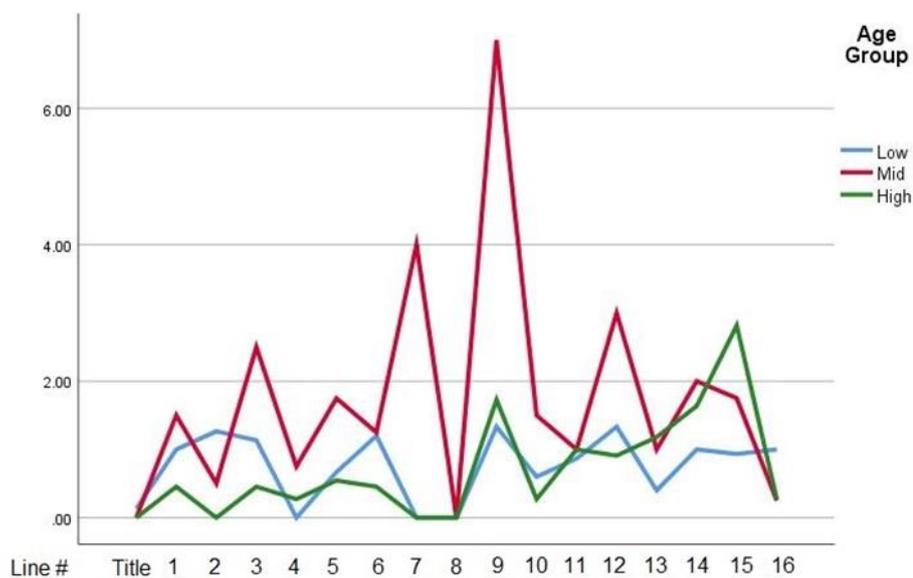
The figures below display the ages of respondents in bar graph form, followed by a table of summary statistics. As we can see from these figures, every undergraduate student fell into the first category (18 – 24), while the other respondents ranged in age from the second (25-34) to the sixth categories (65+). With respect to the category of age, the fact that the youngest group are all students makes it more difficult to interpret the results. Do the results attested by the 18-to-24-year-olds have more to do with their age, or the fact that they all fall into the student category? Any conclusions drawn about these two groups (18-24 vs the rest of the respondents) are thus subject to a degree of uncertainty. Nevertheless, this is a further

opportunity to test a suspicion mentioned in H_2: this is the notion that the "creativity of the reader grows as the institution that controlled it declines" (de Certeau 1984: 172), with 'creativity' here understood as a kind of deviation from institutionally-derived patterns of response.

Fig. 64. Age of Respondents (Survey 1.1)
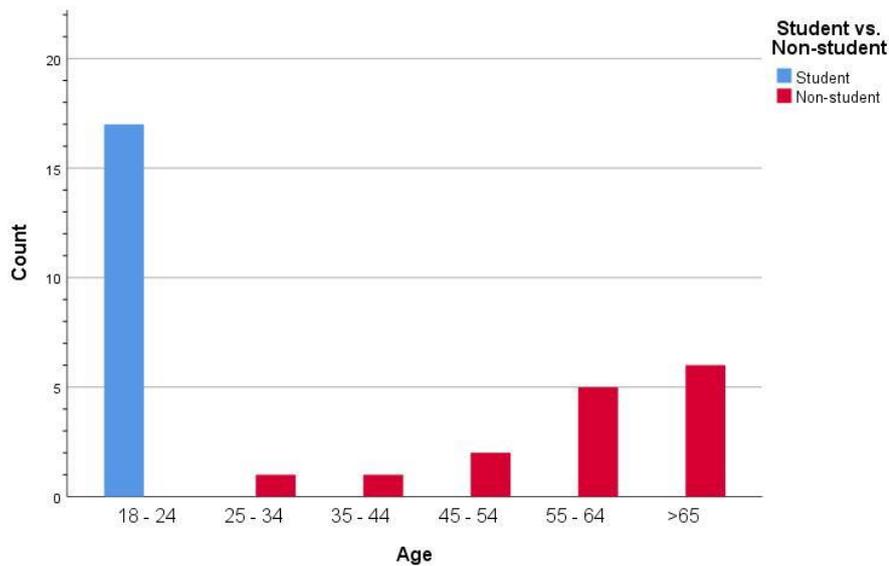
	Student	Non-student
N	17	15
Median group	18 – 24	55 – 64
Mode group	18 – 24	≥65

Fig. 65. Age of Respondents (Survey 1.2)
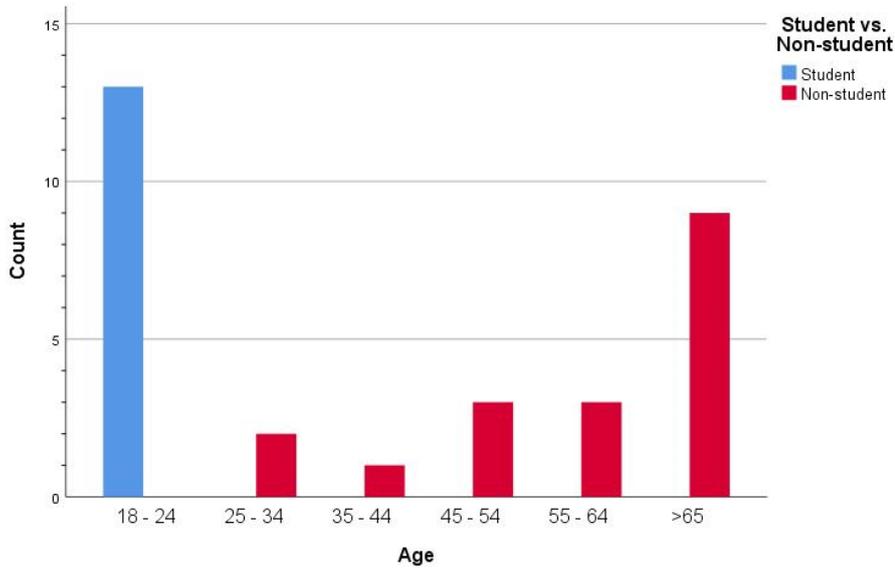
	Student	Non-student
N	13	18
Median group	18 – 24	N/A[a]
Mode group	18 – 24	≥65

[a]There are an even number of groups. In such a case it would normally be possible to calculate an estimated median. Here, however, the two middle groups are 55 – 64 and ≥65. To estimate the median, the class intervals must have equal widths. Here they do not.

To determine whether age and strikingness were correlated, I performed a Spearman rank test. While the results below show there is no correlation between age and number of *items* underlined, it reveals that age and *number of underlinings are indeed negatively correlated* for three poems (Poems A, D and F), with the *p*-value for another poem (Poem H) close to the rejection region.

Fig. 66. Spearman Test: Age vs. Strikingness

	Age vs. No. Words Underlined			Age vs. No. of Underlinings		
	No. Participants (N)	Spearman's rho (r)	p-value	No. Participants (n)	Spearman's rho (r)	p-value
Poem A: Beichman "Shijimi Clams"	32	-0.107	0.560	32	-0.372	0.036
Poem B: Bownas & Thwaite "Two Milliard..."	29	0.171	0.374	29	-0.240	0.210
Poem C: Sato (Midaregami no. 145)	30	0.015	0.938	30	-0.238	0.205
Poem D: Miller & Kudo "Instant"	30	-0.275	0.141	30	-0.412	0.024
Poem E: Rexroth "Shellfish"	30	-0.179	0.343	30	-0.236	0.209
Poem F: Elliott & Kawamura "Two Billion..."	30	-0.133	0.485	30	-0.366	0.047
Poem G: Beichman (Midaregami no. 145)	31	0.010	0.959	31	-0.284	0.122
Poem H: Garza "Shunkan"	30	-0.133	0.482	30	-0.322	0.083

The above results suggest a possible difference in *the manner in which strikingness is perceived*: the older the respondent, the fewer individual instances of strikingness they perceived for these particular poems. This is evidence *against* the 'standard' model of the theory of FG. It also raises further questions of interest: if older respondents made fewer underlinings for these poems, were these less numerous instances cumulatively as striking as the more numerous instances identified by younger respondents? Furthermore, could it be that older respondents tended to underline *more items per individual underlining* than younger respondents? To answer the first question, we might look at whether underlining scores and scores for literariness are associated in any way. (It has been an unspoken assumption in research of this kind that more FG equals more literariness.) In any case, this

question is examined in H$_5$. To answer the second question, we may perform a further Spearman rank test here, this time between age and average number of words underlined per underlining for each respondent.

As the results below show, for Poems B, F and G, there is a moderate positive correlation between age and average length of individual underlining. That is, *the older the respondent, the more likely they were to perceive longer units of text as striking. And by the same token, the younger the respondent, the more likely they were to perceive shorter units of text as striking.* This is evidence against the 'standard' model of the theory of FG. It is true that Poems B and F have some of the longest lines in the text corpora, but their longest lines are only 1 or 2 orthographic words longer than the longest lines in Poems A and E. Moreover, Poem G does not have particularly long lines when compared to the rest of the corpus. And no significant correlation was found in the text with the longest line (Poem C)—that line is about twice as long as the longest line in either Poem B or F. Thus, line length does not seem to be a great explanation for these results.

Fig. 67. Poem B: Age vs. No. Items Underlined Per Underlining

Correlations

			Age	Words Underlined Per Underlining
Spearman's rho	Age	Correlation Coefficient	1.000	.474[*]
		Sig. (2-tailed)	.	.013
		N	27	27
	Words Underlined Per Underlining	Correlation Coefficient	.474[*]	1.000
		Sig. (2-tailed)	.013	.
		N	27	27

*. Correlation is significant at the 0.05 level (2-tailed).

Fig. 68. Poem F: Age vs. Number of Items Underlined Per Underlining

Correlations

			Age	Words Underlined Per Underlining
Spearman's rho	Age	Correlation Coefficient	1.000	.403[*]
		Sig. (2-tailed)	.	.046
		N	25	25
	Words Underlined Per Underlining	Correlation Coefficient	.403[*]	1.000
		Sig. (2-tailed)	.046	.
		N	25	25

*. Correlation is significant at the 0.05 level (2-tailed).

Fig. 69. Poem G: Age vs. Number of Items Underlined Per Underlining

Correlations

			Age	Words Underlined Per Underlining
Spearman's rho	Age	Correlation Coefficient	1.000	.414[*]
		Sig. (2-tailed)	.	.032
		N	27	27
	Words Underlined Per Underlining	Correlation Coefficient	.414[*]	1.000
		Sig. (2-tailed)	.032	.
		N	27	27

*. Correlation is significant at the 0.05 level (2-tailed).

This does not mean that age alone is necessarily responsible for the results. Given that the youngest respondents were all undergraduates, age emerges as a potential measure of the 'grip' of university-influenced patterns of response, with older respondents theoretically less in that grip. Of course, for this to be true, we would have to demonstrate that there *is*

something about modes of reading at the university level that would predispose students to focus on shorter text units. While this is certainly a tall order, it is perhaps not so daunting as it might first appear. Certain possible explanations come to mind. For instance, if literary reading on university courses is geared at least partially toward the production of acceptable interpretations ('acceptable' in Culler's sense of the word), then students might bring certain tendencies to this kind of reading that they would not bring to other reading situations where they are not expected to produce or defend interpretations. Even if students are motivated to study literature at the university level—and are motivated to put in the necessary time to produce 'acceptable' interpretations—they might still feel pressure to work (and read) more quickly than they would under other circumstances. The result of this pressure might be to read with an eye toward identifying *topics* (i.e. 'appropriate' topics to write about) rather than savouring phrases or sentences. In other words, perhaps they are more likely to try to determine the 'aboutness' of the text. This in turn might be motivated by a concern with asking what the text is 'really' about (though, of course, it would remain to be shown whether their literary courses actually emphasised this question). Since there were no undergraduate respondents to Survey 3, where the texts were presented simply as poems (as opposed to translations from Japanese), it is difficult to say what effect the translated-ness of the texts might have had in this regard. However, if students were accustomed to engaging with World Literature in terms of ST reception, they might be less inclined to read for TT manner of expression.

Another possible explanation is suggested by recent research on aging and reading speed. As Gordon, Lowder and Hoedemaker (2016: 167) write in their overview of previous research on the subject: "Older adults generally read more slowly than younger adults, a finding that may be attributable at least partially to a general pattern of age-related cognitive slowing […]. In addition, changes in reading rate among older adults are rooted to some

degree in basic age-related changes in visual perception." The authors note that "[o]lder adults typically make more fixations and more regressive saccades […], and they make longer saccades […]" (ibid.) (Saccades are the "rapid movements" the eyes make "between fixations" [ibid.: 166].) These effects of aging seem to hold across different written languages, too. For instance, in a study of Chinese readers, Wang et al. (2018: 700) report that "[c]ompared to young adults, older adults read more slowly, made more and longer fixations and more regressions, and fixated target words for longer." They write that these "aging effects […] provide further evidence that age-related reading difficulty is found for both alphabetic languages and nonalphabetic languages like Chinese" (ibid.).

But it is not just fixation length that varies with age. As Gordon, Lowder and Hoedemaker (2016: 165) write, "older adults perform as well or better than younger adults when higher-level meanings of a text are assessed." It is in this latter respect that the explanation offered here meshes nicely with the previous one regarding students' potential identification of 'aboutness' at the lexical level. Among other studies, the authors cite the work of Radvansky et al. (2001), whose results showed that "younger adults had better memory than older adults when the recognition task assessed the surface or textbase level; however, older adults outperformed younger adults when the memory tasks assessed aspects of the situation model" (Gordon, Lowder and Hoedemaker 2016: 176). These results support the idea that younger readers may be paying more attention to individual words or short stretches of text. Meanwhile, the fact that older respondents in my surveys tended to identify longer stretches of strikingness for some poems may be related to an age-related tendency to focus on creating situation models of texts. It seems possible that the longer fixation times of older readers not only facilitate the creation of such models, but also have a qualitative effect on the aesthetic experience of the text, insofar as these longer fixations might make longer stretches of text (as opposed to individual words) seem more striking.

While I have focused above on potential differences between younger and older readers, Gordon, Lowder and Hoedemaker (2016: 168) make an interesting observation about the similarities between them: "[O]lder adults show similar (although not identical) effects of word frequency and contextual predictability as do younger readers […]." As we saw in my stylistic analyses, I incorporated both word frequency and contextual predictability in my efforts to predict which parts of the texts would attract the most underlining attention. Indeed, an inspection of the most frequently underlined words for each poems (see tables below) shows that students and non-students—i.e. younger readers and older readers—seemed to largely agree on which words they found striking.

Fig. 70. Poem A: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10	Freq	Line 11
	N/A	N/A	8	dead	2	corner	3	clams	7	alive	3	dawn	12	gobble	3	each	7	cackle	6	evil	1	couldn't	8	half-open
			5	night	2	kitchen	2	little	7	open	2	at	3	i'll	2	and	1	a	6	old	1	had	7	mouth
			4	of	1	of	1	bought	6	mouths			3	up	2	every			6	witch	1	to	4	with
			1	the	1	the	1	evening	4	were			3	you	2	one			3	an			3	all
			1	up			1	i'd											3	like			3	night
			1	woke			1	that															2	sleep
Non-Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10	Freq	Line 11
	N/A	N/A	3	dead	1	a	1	clams	5	mouths	N/A	N/A	4	gobble	N/A	N/A	7	cackle	2	evil	3	had	8	half-open
			3	night	1	corner	1	little	5	open			1	i'll			3	a	2	old	3	to	8	mouth
			3	of	1	in			3	alive			1	up			3	let	2	witch	1	couldn't	5	with
					1	kitchen			2	were			1	you			3	out	1	an	1	help	4	all
					1	of													1	like	1	it	4	night
					1	the																	4	sleep

Fig. 71. Poem B: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8
	2	isolation	8	little	1	and	5	companionship	6	their	2	don't	9	sloop	4	companionship	4	certain
	1	the	6	ball	1	sleeps	4	mars	2	little	2	i	8	wike			2	that's
			4	its	1	wakes	2	with	2	martians	2	know	8	wook				
			3	on	1	works	1	for	1	ball	1	do	6	and				
			1	human					1	on	1	they						
			1	race							1	what						

Non-Students	Freq		Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8
	3	milliard	6	ball	1	and	N/A	N/A	2	ball	2	don't	5	and	1	companionship	N/A	N/A
	2	isolation	6	little	1	sleeps			2	little	2	i	5	sloop	1	earth		
	2	light	3	human	1	wakes			1	martians	2	know	5	wike	1	with		
	2	of	3	its	1	works			1	on	1	do	5	wook				
	2	the	3	on					1	the	1	they	1	maybe				
	2	two	3	race					1	their	1	what	1	they				
	2	years	3	the														

Freq	Line 9	Freq	Line 10	Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16
9	gravitation	13	isolation	1	expands	3	our	2	distends	3	uneasy	4	isolation	8	sneeze
7	universal	11	of			3	unite			1	all	3	milliard	6	involuntary
		7	force			3	wants			1	are	1	involuntary	3	prompts
		5	the			2	all			1	we	1	light	2	an
		3	pulling			2	we					1	two		
		3	together									1	years		
		2	is												

Freq	Line 9	Freq	Line 10	Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16
4	gravitation	8	of	2	expands	4	all	4	distends	2	all	3	light	4	sneeze
3	universal	6	isolation	2	the	4	our	4	universe	2	and	3	milliard	3	involuntary
		6	the	2	universe	4	unite	3	the	2	are	3	years	2	an
		5	force			4	wants			2	so	2	two	2	prompts
		2	is			4	we			2	uneasy	1	isolation		
		2	pulling			2	and			2	we	1	of		
		2	together			2	so					1	the		

Fig. 72. Poem C: Most Frequently Underlined Words (Strikingness)

Students	Freq	Tone-Unit 1	Freq	Tone-Unit 2	Freq	Tone-Unit 3	Freq	Tone-Unit 4	Freq	Tone-Unit 5
	14	merciful	9	traveler	14	love	1	ask	5	me
	5	rainfall			13	my	1	shortcut	4	room
	2	is							4	take
	1	evening							3	a
									3	with

Non-Students	Freq	Tone-Unit 1	Tone	Tone-Unit 2	Tone	Tone-Unit 3	Freq	Tone-Unit 4	Freq	Tone-Unit 5
	9	merciful	4	traveler	2	love	3	ask	5	room
	5	is			2	my	3	don't	4	a
	4	evening					3	shortcut	4	me
	4	rainfall					2	a	4	take
	3	this					2	for	4	with

Fig. 73. Poem D: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
	N/A	N/A	3	dies	10	dies	2	sets	3	dies	4	dies	4	lifts	5	time	1	dies	3	people	7	remorselessly
					8	today					2	today	1	night	4	dies	1	today	2	die		
					6	too					2	too	1	the					2	too		

Non-Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
	N/A	N/A	4	dies	4	dies	N/A	N/A	2	dies	2	dies	4	lifts	4	dies	3	dies	4	die	3	remorselessly
			2	sea	4	today					2	today	4	night	2	times	3	today	3	people		
			2	the	3	too					1	too	4	the					3	too		

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
7	repeatedly	N/A	N/A	1	beauty	N/A	N/A	N/A	N/A	4	remains	4	unfaded	5	undarkened	1	alone	6	ugly	11	dead
3	dying															1	people	1	are	9	drop
																1	then			9	people

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
3	dying	2	again	N/A	N/A	3	dies	2	dies	1	remains	N/A	N/A	N/A	N/A	N/A	N/A	1	ugly	3	dead
3	repeatedly	2	comes			2	sea	1	moon											3	drop
		2	time			1	of													2	people
						1	that														
						1	the														

Fig. 74. Poem E: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7
	1	shellfish	3	midnight	1	bought	6	mouths	7	eat	9	hag's	4	nothing	7	slightly
			2	wake	1	evening	6	slightly	6	will	7	laugh	3	left	6	open
			1	at	1	I	5	open	5	all	2	a	3	night	5	mouth
			1	I	1	last	4	alive	5	breaks			2	of	4	my
					1	little	3	their	5	day			2	the	3	sleep
					1	shellfish	2	are	5	i					3	with
					1	the	2	with	5	them					2	except
									4	when					2	to
Non-Students																
			2	midnight	2	little	9	mouths	2	eat	12	laugh	3	of	9	mouth
			1	at	2	shellfish	9	open	2	I	8	hag's	3	left	8	open
			1	I			9	slightly	2	them	4	a	3	night	8	slightly
			1	wake			6	alive	2	will	1	i	3	nothing	7	my
							5	their	1	all			3	the	3	sleep
							4	with	1	breaks			2	there	2	with
							1	are	1	day			1	afterwards	1	except
									1	when					1	to

Fig. 75. Poem F: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8
	1	billion	4	orb	5	and	4	mars	N/A	N/A	2	martians	6	hararaing	N/A	N/A	N/A	N/A
	1	two	3	small	4	sleep	4	wish			2	orb	4	kiruruing				
			2	beings	4	waken	3	for			2	small	4	neririing				
			2	humans	4	work	3	friends			1	do	4	or				
			2	on	2	sometimes	3	on			1	on						
			2	this							1	their						
											1	what						
Non-Students																		
	N/A	N/A	3	orb	2	sometimes	3	for	2	I've	2	do	4	or	2	Earth	N/A	N/A
			3	small			3	friends	2	no	2	martians	2	hararaing	2	friends		
			2	this			3	mars	2	notion	2	orb	2	kiruruing	2	have		
			1	beings			3	on			2	small	2	neririing	2	like		
			1	human			3	wish			2	their			2	on		
			1	on							2	what			2	sometimes		
											1	on			2	to		
															1	but		
															1	they		

Freq	Line 8	Freq	Line 9	Freq	Line 10	Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16
N/A	N/A	5	solitudes	3	each	4	distorted	4	all	2	expanding	4	all	4	billion	7	sneezed
		4	gravitation	3	other	4	universe	4	we	1	goes	4	uneasy	3	two	4	I
		4	of	3	pulling	3	is	3	another	1	on	4	we	2	of	4	suddenly
		4	power			2	the	3	for	1	the	3	are	2	light-years		
		3	the					3	one	1	universe			2	solitude		
		2	universal					3	seek					1	chill		
		1	is														
N/A	N/A	7	gravitation	3	each	5	distorted	4	another	4	expanding	8	uneasy	9	of	3	sneezed
		7	of	3	other	4	is	4	for	3	goes	6	all	6	chill	1	suddenly
		7	power	3	pulling	4	universe	4	one	3	on	6	are	5	billion		
		7	solitudes			1	because	4	seek	3	universe	6	we	5	light		
		7	universal			1	the	3	all	2	because			5	two		
		6	is					3	we	2	the			5	years		
		6	the											4	solitude		
														3	the		
														1	with		

Fig. 76. Poem G: Most Frequently Underlined Words (Strikingness)

Students	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5
	8	rain	10	desire	3	dear	2	do	5	lodging
	6	the	9	love's	3	traveler	2	not	4	make
	3	falls	6	of			1	shorter	3	here
	1	is					1	way	3	your
	1	tonight							1	but
Non-Students										
	10	rain	10	desire	6	traveler	4	ask	9	lodging
	6	falls	10	love's	3	dear	4	do	6	here
	4	the	6	of			4	not	5	make
	4	tonight					4	shorter	5	your
	4	what					4	the	2	but
	2	is					4	way		

Fig. 77. Poem H: Most Frequently Underlined Words (Strikingness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
	N/A	N/A	5	terminal	6	casualty	4	dusk	6	fatality	4	terminal	2	dawn	4	fatality	6	dead	1	and	N/A	N/A
			3	is	4	a	3	here's	1	a	1	day	1	here's	4	time	3	goes	1	should		
			2	sea	3	day			1	is	1	is			3	a	2	day	1	so		
			1	the	3	is			1	moon					3	is	1	the	1	we		
					3	the																
Non-Students																						
	N/A	N/A	7	terminal	5	casualty	1	here's	5	fatality	2	terminal	N/A	N/A	5	fatality	4	dead	1	and	1	out
			4	is	4	a	1	dusk	3	a	1	is			4	time	3	day	1	should	1	running
			3	sea	3	day			2	is					3	a	3	goes	1	so		
			2	the	3	is			2	moon					3	is	2	the	1	we		
					2	the			2	the												

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
N/A	N/A	6	generous	4	majesty	1	and	3	death	2	vivid	N/A	N/A	N/A	N/A	1	must	2	are	8	death
		5	is	1	if	1	of	1	across									2	few	7	is
		5	itself	1	the	1	moon	1	is									2	ugly	7	need
		5	time			1	sea	1	saved									2	we	7	we
		5	with			1	the											1	the	7	what
																				4	and
2	always	6	generous	1	majesty	1	and	4	death	2	less	N/A	N/A	N/A	N/A	N/A	N/A	4	are	4	death
2	back	5	itself	1	the	1	of	3	across	2	no							4	few	4	is
2	but	5	with			1	moon	3	saved	2	vivid							4	ugly	4	need
2	coming	3	is			1	sea	2	is									3	we	4	we
		3	time			1	the											2	the	4	what
																				2	and

While transliterated Japanese-derived vocabulary was underlined by both younger and older respondents in Poem F (see the mock-Martian words in Line 6), not a single respondent underlined any of the various items printed in Japanese script in Poem H. In fact, this was true for the entire batch of surveys received for Survey 1 and Survey 3, regardless of whether or not any individual survey satisfied the requirements to make it into this analysis. This was true whether items were in *kanji* or *hiragana*; it was true whether items had a transliterated pronunciation gloss; and it was true whether or not there was a gloss for meaning (as there was for the title). Thus, it seems that for the respondents who participated in my survey, the presence of foreign script had the *opposite* effect to what I expected would happen (which is

that respondents might underline at least one or two of these Japanese items).[55] But no: it seemed to preclude entirely the identification of these items as striking. It is possible to interpret this as more evidence that it is not probability of occurrence alone that contributes to an item's strikingness. After all, a keyness analysis comparing Poem H to most any conceivable corpus of English-language poetry would reveal the Japanese items to be highly 'key' to Poem H, as their probability of occurrence would be vanishingly low in the corpus. Indeed, the probability of encountering Japanese script *within* a poem in a representative corpus of *translated* Japanese poetry is itself probably quite low, though such script may indeed appear frequently in paratextual material like book covers or page-design elements. (Rexroth's anthologies of translated Japanese poetry feature the poets' names in calligraphy at the bottom or side of each page.) In using Japanese script within the text of the poem itself, my translation took elements that are more likely to appear in a paratext and put them in the actual text. And yet even this deviation was not enough to make the items seem striking to readers. Perhaps the presence of largely unintelligible text shunted readers' attention toward those parts of the text they *could* read, with the Japanese script itself becoming a kind of new BG. If this is the case, then perhaps intelligibility is what prevents stretches of text from fading immediately into the BG; *thus, strikingness in poetry may depend in part on intelligibility*. One respondent arguably invoked this category to explain why they didn't respond to the Likert scales for Poem H: "I cannot engage with this piece in spite of a slight familiarity with Japanese, so decline to respond." This is an interesting response insofar as it seems to suggest that only readers who can read in both languages should fill out the survey for this poem. But it may also suggest a belief that the two scripts (Japanese and English) work together to comprise an aesthetic whole—and that special knowledge is required to

---

[55] I recall what happened when I distributed surveys at one venue: one respondent flipped immediately to Poem H and exclaimed, "We even get to learn some Japanese!" Given this kind of surprised reaction, it came as something of a surprise when no one singled out those same Japanese items as striking.

adequately evaluate that aesthetic whole. So on the one hand, this response can be read as a modest demurral born out of respect for a holistic experience that this respondent presumes is available only to bilingual readers; and on the other hand, it is a curiously self-denying response: why not take whatever aesthetic enjoyment was to be found, considering that other respondents surveyed in this venue presumably would? In fact, another respondent indicated they were not sure whether the glosses "were part of the poem or an additional clarification of the Japanese symbols for the people who can read Japanese." Nevertheless, they scored the poem quite highly for both literariness and enjoyment. Implicated in this question is a presumed lack of access to the meaning of most of the items in Japanese script—while some items had glosses for pronunciation, this was not enough to tempt the first reader to evaluate the poem. Van Peer's (1986: 68) intuition that "people generally will tend to look for meaning in reading first, and will give secondary attention to formal or phonological structures" thus finds some support here.

Van Peer's intuition seems true for at least one other reader: one respondent circled the item "Milliard" in the title of Poem B and scribbled a question mark next to it, presumably to indicate that they were unfamiliar with the word. This response suggests that perhaps not all un-intelligibilities are created equal. As this respondent was replying to a copy of the survey with Poems A, B, C and D, they did not encounter Poem H (with the Japanese script items). However, it is notable that the only instance of a word being singled out this way—out of all the surveys returned for Surveys 1, 2 and 3—should be a British-English item and not an item from the Japanese. 'Milliard' is an uncommon word these days, but its phonotactics and spelling (like the phonotactics and spelling of the mock-Martian words) would presumably mark it out as at least a nonce-word (and therefore a possible word) in English. Thus, if we were to try to create a typology of strikingness in future, this fact would surely be a part.

But to return to the question of Poem H, the lack of attested strikingness might also be

due in part to fatigue associated with the ubiquity of Japanese and Chinese as design elements in Anglophone contexts, e.g., tattoos, clothing brand graphic design—I refer here to the UK clothing brand Superdry, whose labels bear the strange imperative "極度乾燥（しなさい）" (please dry to the maximum degree).

Thus, while H_4 is explicitly concerned with the theory of FG, the above discussion has interesting implications for the theory of foreignisation. One is that strict adherence to SL structures—even to the point of retaining orthography—may not stand out to readers. Another is that, as we will see in our discussion of enjoyment, it is nevertheless still possible for some effectively monolingual readers to enjoy bilingual writing.

One final thing to note about the lists of most frequently underlined words: words that had been demoted from FG inventories due to their association with title words nevertheless appear in the lists of 'Most Frequently Underlined Words' for every stylistically analysed poem, apart from Sato's *Midaregami* no. 145. Thus, it seems it was not necessary to demote such items in the first place. *Association with title words does not emerge as a relevant factor in respondents' underlining behaviour on the whole.*

Of course, it is possible for these two groups of readers to agree on particular *words*, while nevertheless disagreeing in terms of which *lines* attracted the most or least underlining attention. We can see this from the plots below, which show the average number of items underlined per group on a line-by-line basis. Trend lines that cross each other suggest qualitative differences in the way different groups responded to the texts.

However, this is complicated by the fact that older readers skip more words (Gordon, Lowder and Hoedemaker 2016: 169). These authors cite research showing that older readers have a "tendency to show larger and more consistent effects of word frequency on word-skipping rates […]" (ibid.), which "suggests that they tend to rely on (partly) visual and word frequency information to 'guess' the identity of upcoming words, thereby skipping more

307

words in an attempt to speed up reading rate" (ibid.). Nevertheless, it is another potential starting point for future research. (And another potential thing to study: does foreknowledge of the translated-ness of texts lead to longer reading times? Perhaps some readers assume that because a text is translated it may take more mental effort to read, and therefore they expend more effort doing so.)

Fig. 78. Poem A: Average No. Words Underlined Per Line, Per Group (Strikingness)
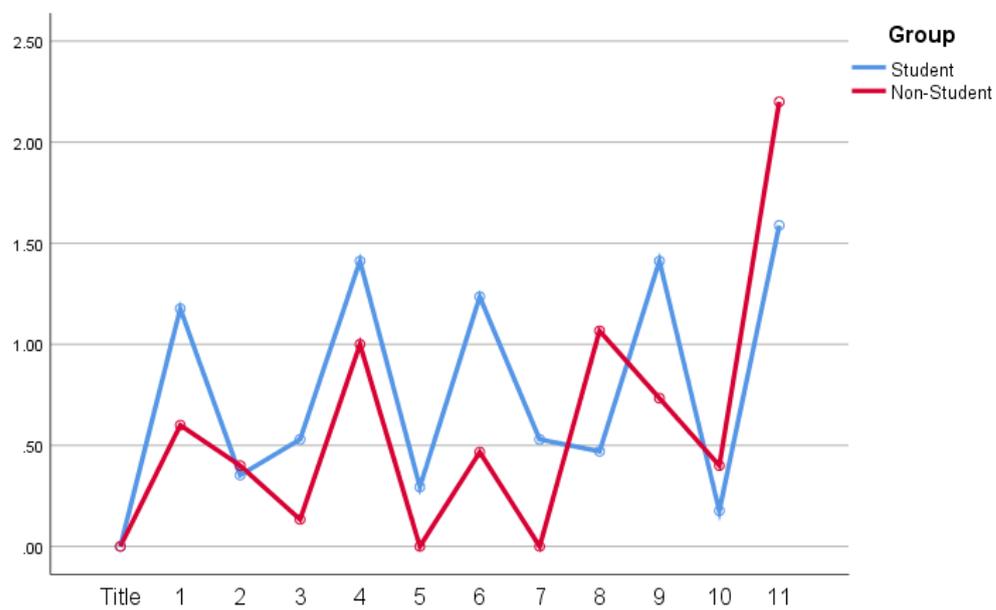
Fig. 79. Poem B: Average Number of Words Underlined Per Line, Per Group (Strikingness)
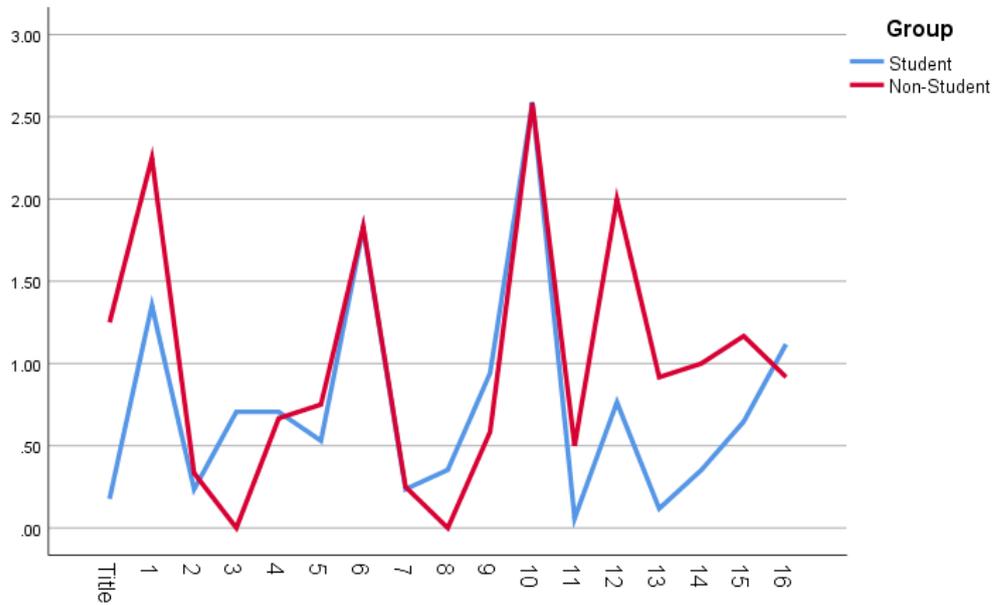
Fig. 80. Poem C: Average Number of Words Underlined Per Tone-Unit, Per Group (Strikingness)
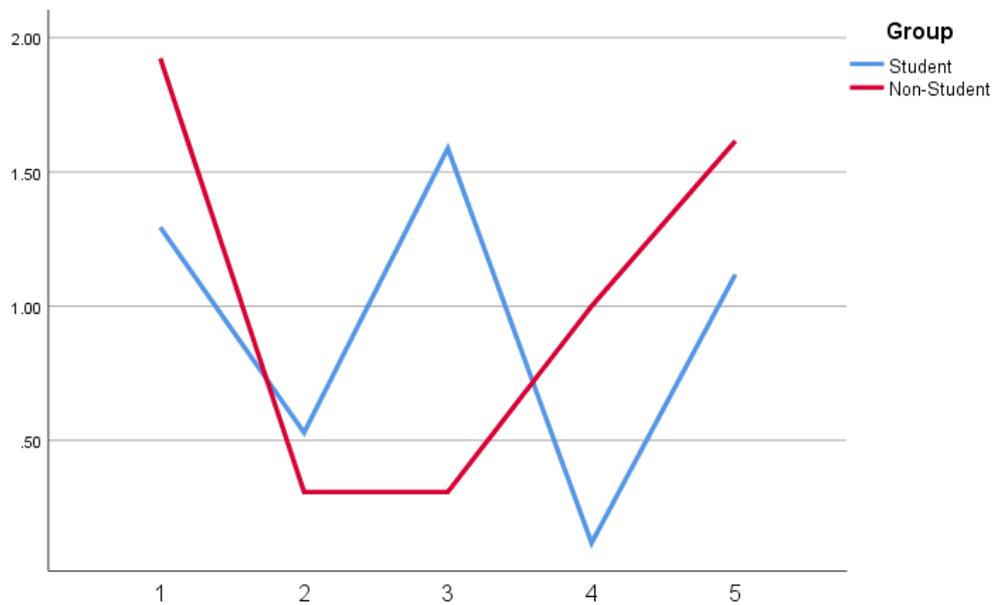
Fig. 81. Poem D: Average Number of Words Underlined Per Line, Per Group (Strikingness)
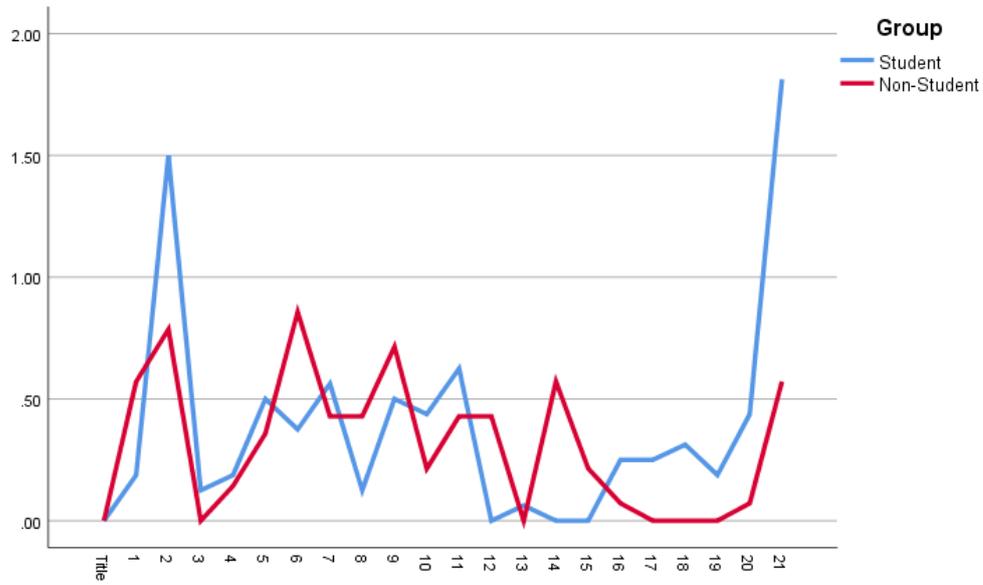
Fig. 82. Poem E: Average Number of Words Underlined Per Line, Per Group (Strikingness)
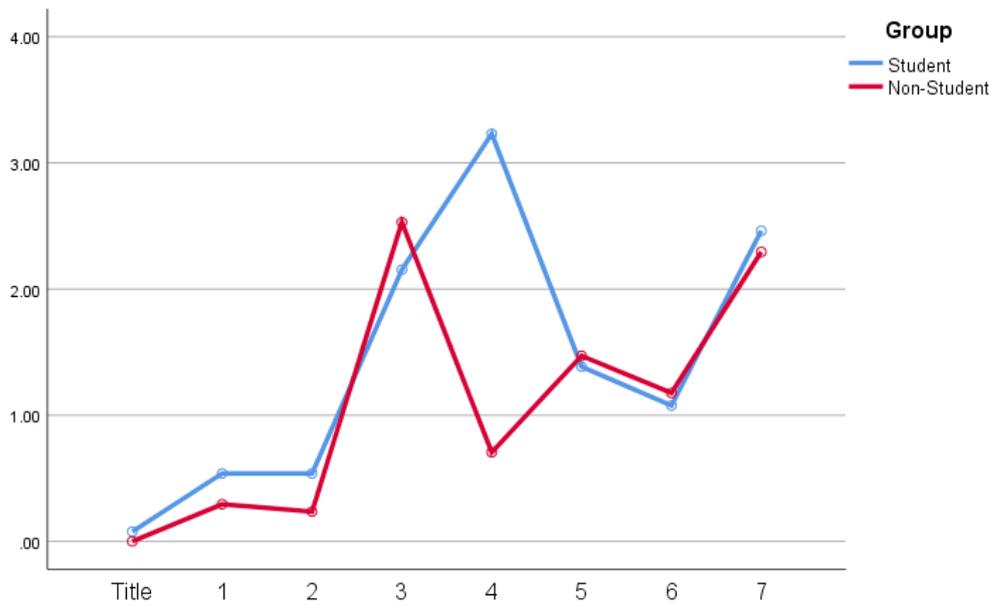
Fig. 83. Poem F: Average Number of Words Underlined Per Line, Per Group (Strikingness)
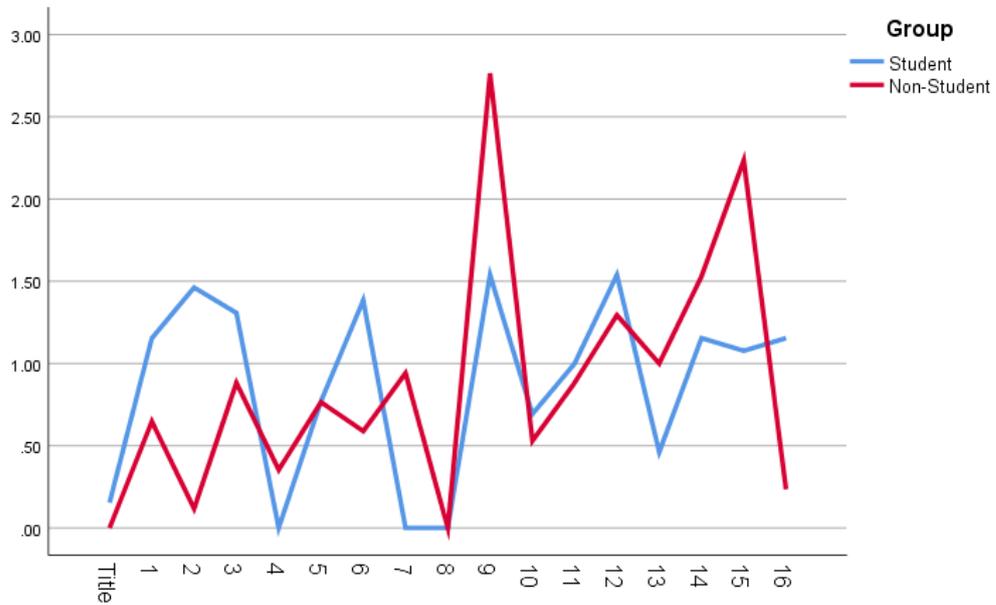
Fig. 84. Poem G: Average Number of Words Underlined Per Line, Per Group (Strikingness)
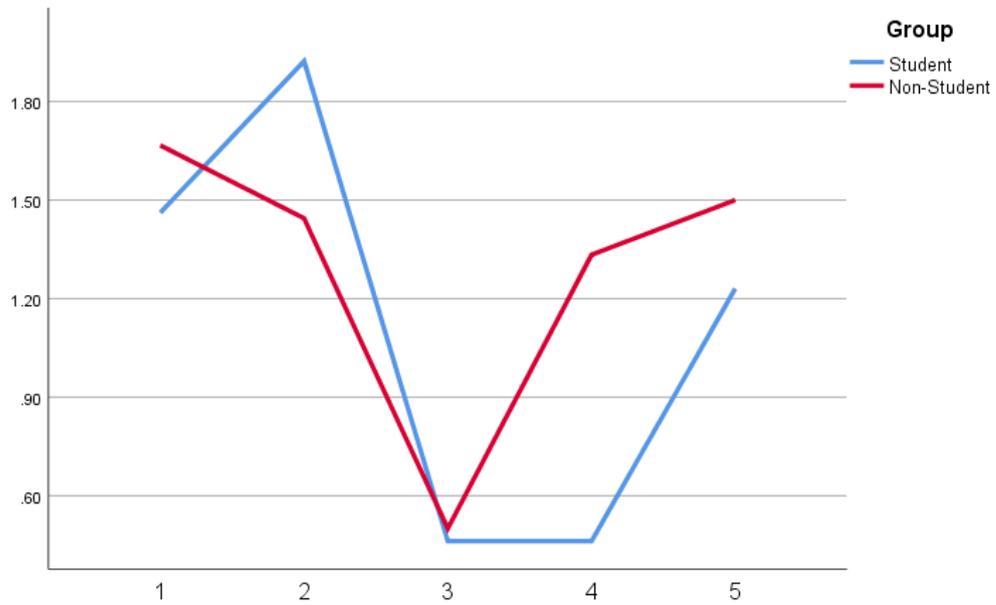
Fig. 85. Poem H: Average Number of Words Underlined Per Line, Per Group (Strikingness)
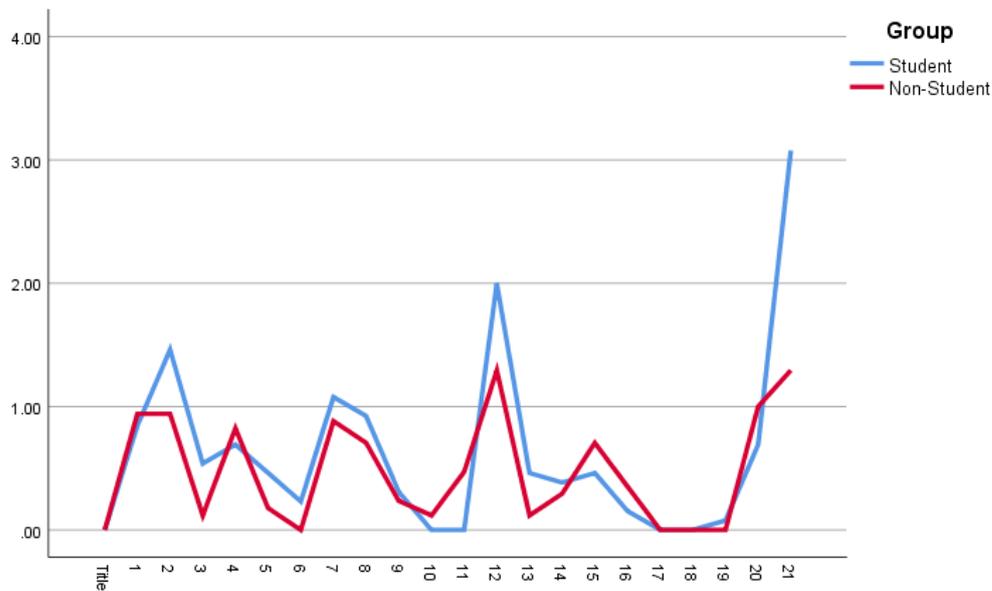
For some poems, for instance Poem E, the trajectories of the trend lines seem fairly

similar—where the trend line for one group rises, so does the line for the other group, and

where one falls, so does the other. However, even in cases like this, there is always at least

one instance where the trend lines take different directions (for Poem E, it is between verse

lines 3 and 4). Meanwhile, for some poems, the trend line for one group is convex while for

the other group it is concave at the same verse line (see e.g. verse line 2 in Poem F). In

general, it seems that the longer the poem, the more instances of this kind of divergence we

can see throughout the text. For poems where this is the case, we might say that the two

groups had two qualitatively different experiences of the text (at least in terms of striking

verse lines). If we can conceive of the reading of a given poem as a kind of journey with

peaks and troughs, then for the longer poems especially, these peaks and troughs came at

different points, depending on which group (student or non-student) a respondent belonged

to. The main exception to this is Poem H, which consists of 22 lines (including title), and yet

which shows a surprising degree of qualitative agreement between the responses of the two

groups. As we can see from the graph for Poem H, the two groups seemed to agree on the relative strikingness of many lines. While this is rather puzzling, one potential explanation does come to mind: perhaps the disruption created by the Japanese items in the beginning of every second and third line in the staggered triplets acted as a kind of 'reset,' interfering with how the lines in English interacted with each other. That is, perhaps breaking up the English text in this manner 'reset' readers' expectations concerning lexical probabilities, so that readers perceived the text in a more fragmentary way. Thus, there was less chance of an early perceptual difference 'snowballing' for one group such that it affected subsequent expectations surrounding lexical probabilities. More research would of course be required to test this idea. However, we can at least put forward a falsifiable hypothesis here. In any case, all in all, results are again mixed.

5.2.5: HYPOTHESIS 5 (PERCEPTION OF LITERARINESS)

H_5 concerns the relationship between FG and literariness. H_a predicted that the more foregrounding readers indicated, the higher they would rate the text in terms of literariness. H_0 stated that the amount of perceived FG would have no significant association with ratings for literariness. The test instruments used to collect data were the underlining test for strikingness (from H_1), as well as the Likert-scaled items concerning literariness. A significant positive correlation on a Spearman test would suggest that the more FG respondents underlined, the more 'literary' the particular text seemed. In this case, I performed multiple Spearman rank tests. First, I checked whether the rankings produced by the data from students and non-students agreed with each other. This would tell me whether they tended to consider the same poems the most 'literary.' Next, I checked for correlations between the total number of FG features (as identified by stylistic analysis) and literariness scores per poem. I included this additional test because the relative amount of FG perceived in a poem might not be accurately reflected in the number of words underlined, and it might be the case that the total amount of FG (as identified in stylistic analysis) is a better predictor of scores for impression of literariness. Finally, I checked for correlations between the number of words underlined and the literariness scores for each poem.

Results for Survey 1.1 and Survey 1.2 show:

- **there was a very strong positive correlation between rankings of poems produced by student and non-student data for median number of words underlined per respondent** (according to a Spearman rank test)
    - *this is evidence **in favour of** the theory of FG*
- **however, the above result may be explained by the fact that the longer the poem, the more items respondents tended to underline**
    - *this is evidence **against** the theory of FG*

314

- **moreover, there was no correlation between rankings of poems produced by student and non-student data for median literariness scores** (according to a Spearman rank test)
  - *this is evidence **against** the theory of FG*
- **moreover, there was no correlation between amount of FG (as identified by stylistic analysis) and median literariness scores** (according to a Spearman rank test)
  - *this is evidence **against** the theory of FG*
- **for Poems A, B, C, E, G and H, there were no significant correlations between the number of underlined words and literariness scores**, for student respondents (according to the Spearman rank test)
  - *this is evidence **against** the theory of FG*
- **for Poems D, E and G, there were no significant correlations between the number of underlined words and literariness scores**, for non-student respondents (according to the Spearman rank test)
  - *this is evidence **against** the theory of FG*
- **for Poems B, C, D, F and H, there were significant *negative* correlations between number of underlined words and literariness scores**, depending on the respondent group (according to the Spearman rank test)
  - *this is evidence **against** the theory of FG*

All in all, we have found some good evidence *against* the theory of foregrounding. While students and non-students both tended to underline more words in poems for which I had identified larger amounts of FG, there was no correlation between the amount of FG I had identified and median literariness scores for the poems. *Thus, it seems that the amount of*

FG the researcher identifies by stylistic analysis has no bearing on whether respondents will consider a poem to be literary. There *was* a correlation between the amount of FG I had identified and the number of items respondents tended to underline per poem. However, this could also be explained by the text length—the longer the poem, the more respondents tended to underline.

The rationale behind this test was to investigate the assumption that FG contributes in a necessarily positive way to the aesthetic experience of texts—including impression of literariness. Such an issue should be of interest to TS researchers for a couple of reasons. First, the degree to which a text resembles other texts previously encountered by readers as literature would seem to depend on the particular literary system under discussion. As I tried to show in section 2.1.1, such 'literariness' may depend upon factors *other* than features associated with defamiliarisation (e.g. the diglossic condition of Japanese during the Meiji period). Secondly, if the responses to Venuti's own foreignising translations are any indication, then translations may indeed foreground precisely those features which are *unlike* what readers have previously encountered as literature.

The tables below show both the mean and median number of words underlined per respondent per poem. While the distribution of this variable of interest passes normality tests, the use of means as a measure of central tendency in cases like this should be considered carefully. As Ghasemi and Zahediasl (2012: 487) write: "For small sample sizes, normality tests have little power to reject the null hypothesis and therefore small samples most often pass normality tests." Thus, in certain instances, the median may be a more robust measure of central tendency. I have also included SDs for reference. Where literariness scores are concerned, the scale used to collect data was ordinal, so median values are arguably the most appropriate measure of central tendency. Literariness scores were calculated for each respondent by summing their responses to two questions on the Likert scale: 1) Is this poem

an example of good literature? and 2) In your opinion, how literary is this poem? I decided to sum these two items as they were the ones which related most to the construct of literariness I had in mind. Specifically, I was interested in how similar a given text seemed to other texts that respondents had encountered before as belonging to the category of literature, however they might define that category.

Fig. 86. No. Words Underlined vs. Literariness Scores (Students)

	Poem A	Poem B	Poem C	Poem D	Poem E	Poem F	Poem G	Poem H
Mean Words Underlined per Respondent	9.18	12.71	4.65	8.44	11.46	14.85	5.54	13.85
Median Words Underlined per Respondent	7.00	11.00	4.00	8.50	10.00	16.00	5.00	14.00
SD (Sample)	6.10	6.54	2.42	3.41	7.43	10.26	3.04	10.55
SD (Population)	5.91	6.34	2.35	3.30	7.14	9.86	2.93	10.14
Mean Literariness Score	7.82	10.12	7.71	9.44	9.15	10.46	8.31	10.38
Median Literariness Score	9.00	10.00	8.00	9.50	10.00	10.00	8.00	10.00
SD (Sample)	2.53	1.32	2.69	2.34	3.16	1.81	2.32	1.98
SD (Population)	2.46	1.28	2.61	2.26	3.03	1.74	2.23	1.90

Fig. 87. No. Words Underlined vs. Literariness Scores (Non-Students)

	Poem A	Poem B	Poem C	Poem D	Poem E	Poem F	Poem G	Poem H
Mean Words Underlined per Respondent	7.00	17.00	5.15	6.86	8.71	14.76	6.44	10.47
Median Words Underlined per Respondent	8.00	15.00	4.00	5.50	9.00	13.00	6.50	10.00
SD (Sample)	4.84	12.88	4.83	5.64	6.38	13.09	4.37	8.99
SD (Population)	4.68	12.34	4.64	5.44	6.19	12.70	4.25	8.73
Mean Literariness Score	6.60	8.42	8.31	7.93	9.29	8.76	10.33	7.12
Median Literariness Score	6.00	9.00	8.00	8.00	10.00	9.00	11.00	8.00
SD (Sample)	2.77	3.78	2.59	3.02	2.59	2.97	2.70	3.46
SD (Population)	2.68	3.62	2.49	2.91	2.52	2.88	2.62	3.36

When we compare these tables, we notice that the data from the two groups produce different rankings of the poems. This is true whether we rank the poems in terms of median words underlined per respondent, or in terms of median literariness scores. However, there seems to be less agreement when the poems are ranked by median literariness scores. In short, it does not look like the two groups would agree with each other about which texts seemed the most literary (i.e. when the poems are ranked by median literariness scores).

Fig. 88. Poems Ranked by Median Literariness Score (Students)

Poem	Mean Rank	Median Lit Score
B	2.5	10
E	2.5	10
F	2.5	10
H	2.5	10
D	5	9.5
A	6	9
C	7.5	8
G	7.5	8

Fig. 89. Poems Ranked by Median Literariness Score (Non-Students)

Poem	Mean Rank	Median Lit Score
G	1	11
E	2	10
B	3.5	9
F	3.5	9
C	6	8
D	6	8
H	6	8
A	8	6

The results from a Spearman rank test corroborate our suspicions. As we can see from the tables below, there is a strong positive correlation between the student and non-student groups' rankings of the poems in terms of *median number of words underlined*. Both groups tended to underline the most words in the same poems. Thus, responses on this measure seem more text-directed than reader-directed. As this supports the theory of FG, I have highlighted the relevant *p*-value in green. However, a Spearman rank test comparing rankings by median literariness scores found no correlation between the student and non-student groups. The two groups did not tend to find the same poems to be the most literary. Thus, responses on this particular measure seem more reader-directed than text-directed. Since this does not support the theory of FG, I have highlighted the relevant *p*-value in yellow.

Fig. 90. Spearman Test: Poems Ranked by Median No. Words Underlined

			Poems Ranked by Median Number of Words Underlined (Students)	Poems Ranked by Median Number of Words Underlined (Non-Students)
Spearman's rho	Poems Ranked by Median Number of Words Underlined (Students)	Correlation Coefficient	1.000	.857**
		Sig. (2-tailed)	.	.007
		N	8	8
	Poems Ranked by Median Number of Words Underlined (Non-Students)	Correlation Coefficient	.857**	1.000
		Sig. (2-tailed)	.007	.
		N	8	8

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 91. Spearman Test: Poems Ranked by Median Literariness Scores

			Poems Ranked by Median Lit Score (Students)	Poems Ranked by Median Lit Score (Non-Students)
Spearman's rho	Poems Ranked by Median Lit Score (Students)	Correlation Coefficient	1.000	.158
		Sig. (2-tailed)	.	.709
		N	8	8
	Poems Ranked by Median Lit Score (Non-Students)	Correlation Coefficient	.158	1.000
		Sig. (2-tailed)	.709	.
		N	8	8

However, there is more to say about the positive correlation between the student and non-student groups regarding the median number of underlined words. As I noted earlier, the number of items underlined for strikingness may not accurately reflect the *amount of FG* in a given text (as identified by my earlier stylistics analyses). Thus, I performed an additional Spearman rank test to determine whether *amount of FG* (as identified by my analyses) was correlated with the number of items underlined per poem. As the results below show, there was indeed a strong correlation. The poems for which I had identified the most FG also

attracted the most underlining activity. Insofar as this matches the predictions of the theory, this finding supports the theory of FG. Thus, I have highlighted this result in green.

Fig. 92. Spearman Test: FG Features (Stylistic Analysis) vs. No. Items Underlined

			Mean Rank by FG Features	Mean Rank by Median Number of Items Underlined
Spearman's rho	Mean Rank by FG Features	Correlation Coefficient	1.000	.829*
		Sig. (2-tailed)	.	.042
		N	6	6
	Mean Rank by Median Number of Items Underlined	Correlation Coefficient	.829*	1.000
		Sig. (2-tailed)	.042	.
		N	6	6

*. Correlation is significant at the 0.05 level (2-tailed).

However, when we look back at the tables at the start of this section, we see that the poems with the most items underlined seem to be the longest poems. This seems common-sensical: the longer the poem, the more opportunities for underlining! Thus, poem length might predict the number of words underlined just as well as my stylistic analyses. To test this, I ranked poems by length, and then again by number of FG features identified in stylistic analysis, and ran another Spearman test.[56] The results show that the longer texts indeed had more FG features identified by my analysis. Since this is a competing explanation for why respondents underlined more items in poems with more FG, it does not support the theory of FG, and I have highlighted the result in yellow.

---

[56] Since no stylistic analyses were performed on Poems D and H, they had to be excluded from this test.

Fig. 93. Spearman Test: FG Features (Stylistic Analysis) vs. Poem Length

			Mean Rank by FG Features	Mean Rank by Orthographic Length
Spearman's rho	Mean Rank by FG Features	Correlation Coefficient	1.000	.943**
		Sig. (2-tailed)	.	.005
		N	6	6
	Mean Rank by Orthographic Length	Correlation Coefficient	.943**	1.000
		Sig. (2-tailed)	.005	.
		N	6	6

**. Correlation is significant at the 0.01 level (2-tailed).

So: how can we disentangle the effects of text length from the total number of FG features identified by stylistic analysis? One way would be to perform a multiple regression analysis; however, that would be a parametric test, and as Pett (1997: 278) notes, there is "no nonparametric equivalent to the […] multiple regression." Otherwise, it seems we would need to design another survey, such that it included some long texts with comparatively few FG features and some short texts with a relative abundance of the same. Nevertheless, at this stage, we cannot definitively say that the *amount of FG* in a poem is the reason why some poems had more items underlined that others. It could just as well be that the longer the poem is, the more words respondents are likely to underline.

But what if *amount of FG* (as identified by my analyses) was correlated with overall median literariness scores? To answer this question, I performed yet another test, but as we can see below, no correlation was found. In other words, it did not seem to matter how much FG I had identified—this had no real bearing on whether respondents considered a particular text to be literary. Since this is evidence against the theory of FG, I have highlighted the relevant *p*-value in yellow.

Fig. 94. Spearman Test: FG Features (Stylistic Analysis) vs. Median Literariness Scores

			Mean Rank by FG Features	Mean Rank by Median Lit Scores
Spearman's rho	Mean Rank by FG Features	Correlation Coefficient	1.000	.031
		Sig. (2-tailed)	.	.954
		N	6	6
	Mean Rank by Median Lit Scores	Correlation Coefficient	.031	1.000
		Sig. (2-tailed)	.954	.
		N	6	6

Since D and H were left out of the above tests, I thought it prudent to perform one last test—a test which could include those poems. Also, the criticism could be levied that results might be different if I had chosen other Likert scale items on which to base literariness scores. Thus, I performed some final Spearman rank tests to see whether underlining in the poems was correlated in any way with any other Likert questions which could conceivably be considered to measure impression of literariness. This time, I accounted not only for number of words underlined, but number of individual underlinings as well. Significant results in favour of the theory of FG are highlighted in green. Significant results that show the *opposite* to that predicted by the theory are highlighted in red.

Fig. 95.

Poem A: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.415	-.232	-.211	-.202	-.169
		Sig. (2-tailed)	.	.098	.371	.416	.437	.518
		N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.415	1.000	-.226	-.040	-.265	.190
		Sig. (2-tailed)	.098	.	.384	.878	.305	.464
		N	17	17	17	17	17	17
	Dixon Scale	Correlation Coefficient	-.232	-.226	1.000	.754**	.715**	.712**
		Sig. (2-tailed)	.371	.384	.	.000	.001	.001
		N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.211	-.040	.754**	1.000	.942**	.917**
		Sig. (2-tailed)	.416	.878	.000	.	.000	.000
		N	17	17	17	17	17	17
	Ex of Good Lit?	Correlation Coefficient	-.202	-.265	.715**	.942**	1.000	.746**
		Sig. (2-tailed)	.437	.305	.001	.000	.	.001
		N	17	17	17	17	17	17
	How Literary Is This Poem?	Correlation Coefficient	-.169	.190	.712**	.917**	.746**	1.000
		Sig. (2-tailed)	.518	.464	.001	.000	.001	.
		N	17	17	17	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 96.

Poem A: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.820**	.003	-.100	-.022	-.148
		Sig. (2-tailed)	.	.000	.992	.724	.937	.599
		N	15	15	15	15	15	15
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.820**	1.000	-.391	-.445	-.327	-.451
		Sig. (2-tailed)	.000	.	.149	.096	.234	.091
		N	15	15	15	15	15	15
	Dixon Scale	Correlation Coefficient	.003	-.391	1.000	.920**	.920**	.764**
		Sig. (2-tailed)	.992	.149	.	.000	.000	.001
		N	15	15	15	15	15	15
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.100	-.445	.920**	1.000	.935**	.909**
		Sig. (2-tailed)	.724	.096	.000	.	.000	.000
		N	15	15	15	15	15	15
	Ex of Good Lit?	Correlation Coefficient	-.022	-.327	.920**	.935**	1.000	.718**
		Sig. (2-tailed)	.937	.234	.000	.000	.	.003
		N	15	15	15	15	15	15
	How Literary Is This Poem?	Correlation Coefficient	-.148	-.451	.764**	.909**	.718**	1.000
		Sig. (2-tailed)	.599	.091	.001	.000	.003	.
		N	15	15	15	15	15	15

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 97

Poem B: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.142	-.161	-.156	-.114	-.198
		Sig. (2-tailed)	.	.588	.536	.550	.664	.447
		N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.142	1.000	.065	.188	.017	.215
		Sig. (2-tailed)	.588	.	.803	.470	.949	.407
		N	17	17	17	17	17	17
	Dixon Scale	Correlation Coefficient	-.161	.065	1.000	.531*	.722**	.291
		Sig. (2-tailed)	.536	.803	.	.028	.001	.257
		N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.156	.188	.531*	1.000	.868**	.891**
		Sig. (2-tailed)	.550	.470	.028	.	.000	.000
		N	17	17	17	17	17	17
	Ex of Good Lit?	Correlation Coefficient	-.114	.017	.722**	.868**	1.000	.570*
		Sig. (2-tailed)	.664	.949	.001	.000	.	.017
		N	17	17	17	17	17	17
	How Literary Is This Poem?	Correlation Coefficient	-.198	.215	.291	.891**	.570*	1.000
		Sig. (2-tailed)	.447	.407	.257	.000	.017	.
		N	17	17	17	17	17	17

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 98

Poem B: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.931**	-.565	-.674*	-.528	-.729**
		Sig. (2-tailed)	.	.000	.056	.016	.078	.007
		N	12	12	12	12	12	12
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.931**	1.000	-.594*	-.621*	-.517	-.654*
		Sig. (2-tailed)	.000	.	.042	.031	.085	.021
		N	12	12	12	12	12	12
	Dixon Scale	Correlation Coefficient	-.565	-.594*	1.000	.947**	.980**	.746**
		Sig. (2-tailed)	.056	.042	.	.000	.000	.005
		N	12	12	12	12	12	12
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.674*	-.621*	.947**	1.000	.948**	.893**
		Sig. (2-tailed)	.016	.031	.000	.	.000	.000
		N	12	12	12	12	12	12
	Ex of Good Lit?	Correlation Coefficient	-.528	-.517	.980**	.948**	1.000	.733**
		Sig. (2-tailed)	.078	.085	.000	.000	.	.007
		N	12	12	12	12	12	12
	How Literary Is This Poem?	Correlation Coefficient	-.729**	-.654*	.746**	.893**	.733**	1.000
		Sig. (2-tailed)	.007	.021	.005	.000	.007	.
		N	12	12	12	12	12	12

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

324

Fig. 99

Poem C: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

		No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks) — Correlation Coefficient	1.000	.767**	-.182	-.318	-.334	-.283
	Sig. (2-tailed)	.	.000	.484	.213	.190	.270
	N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks) — Correlation Coefficient	.767**	1.000	-.049	-.332	-.259	-.377
	Sig. (2-tailed)	.000	.	.851	.193	.316	.136
	N	17	17	17	17	17	17
	Dixon Scale — Correlation Coefficient	-.182	-.049	1.000	.817**	.878**	.648**
	Sig. (2-tailed)	.484	.851	.	.000	.000	.005
	N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary? — Correlation Coefficient	-.318	-.332	.817**	1.000	.941**	.951**
	Sig. (2-tailed)	.213	.193	.000	.	.000	.000
	N	17	17	17	17	17	17
	Ex of Good Lit? — Correlation Coefficient	-.334	-.259	.878**	.941**	1.000	.804**
	Sig. (2-tailed)	.190	.316	.000	.000	.	.000
	N	17	17	17	17	17	17
	How Literary Is This Poem? — Correlation Coefficient	-.283	-.377	.648**	.951**	.804**	1.000
	Sig. (2-tailed)	.270	.136	.005	.000	.000	.
	N	17	17	17	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 100

Poem C: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

		No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks) — Correlation Coefficient	1.000	.867**	-.515	-.520	-.577*	-.471
	Sig. (2-tailed)	.	.000	.072	.069	.039	.104
	N	13	13	13	13	13	13
	No. of Underlinings (Mean Ranks) — Correlation Coefficient	.867**	1.000	-.119	-.139	-.222	-.101
	Sig. (2-tailed)	.000	.	.699	.650	.467	.741
	N	13	13	13	13	13	13
	Dixon Scale — Correlation Coefficient	-.515	-.119	1.000	.765**	.741**	.672*
	Sig. (2-tailed)	.072	.699	.	.002	.004	.012
	N	13	13	13	13	13	13
	Ex. of Good Lit? + How Literary? — Correlation Coefficient	-.520	-.139	.765**	1.000	.946**	.959**
	Sig. (2-tailed)	.069	.650	.002	.	.000	.000
	N	13	13	13	13	13	13
	Ex of Good Lit? — Correlation Coefficient	-.577*	-.222	.741**	.946**	1.000	.860**
	Sig. (2-tailed)	.039	.467	.004	.000	.	.000
	N	13	13	13	13	13	13
	How Literary Is This Poem? — Correlation Coefficient	-.471	-.101	.672*	.959**	.860**	1.000
	Sig. (2-tailed)	.104	.741	.012	.000	.000	.
	N	13	13	13	13	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Fig. 101. Poem D: Spearman Rank Correlations—Number of Words Underlined vs. Literariness Scores (Student Respondents)

Correlations

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.512*	-.568*	-.543*	-.390	-.641**
		Sig. (2-tailed)	.	.042	.022	.030	.135	.007
		N	16	16	16	16	16	16
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.512*	1.000	-.251	-.183	-.135	-.201
		Sig. (2-tailed)	.042	.	.348	.499	.618	.456
		N	16	16	16	16	16	16
	Dixon Scale	Correlation Coefficient	-.568*	-.251	1.000	.901**	.887**	.808**
		Sig. (2-tailed)	.022	.348	.	.000	.000	.000
		N	16	16	16	16	16	16
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.543*	-.183	.901**	1.000	.943**	.948**
		Sig. (2-tailed)	.030	.499	.000	.	.000	.000
		N	16	16	16	16	16	16
	Ex of Good Lit?	Correlation Coefficient	-.390	-.135	.887**	.943**	1.000	.804**
		Sig. (2-tailed)	.135	.618	.000	.000	.	.000
		N	16	16	16	16	16	16
	How Literary Is This Poem?	Correlation Coefficient	-.641**	-.201	.808**	.948**	.804**	1.000
		Sig. (2-tailed)	.007	.456	.000	.000	.000	.
		N	16	16	16	16	16	16

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 102

Poem D: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.882**	.000	.097	.080	.241
		Sig. (2-tailed)	.	.000	1.000	.742	.787	.407
		N	14	14	14	14	14	14
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.882**	1.000	.314	.426	.405	.513
		Sig. (2-tailed)	.000	.	.275	.129	.151	.061
		N	14	14	14	14	14	14
	Dixon Scale	Correlation Coefficient	.000	.314	1.000	.905**	.947**	.776**
		Sig. (2-tailed)	1.000	.275	.	.000	.000	.001
		N	14	14	14	14	14	14
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	.097	.426	.905**	1.000	.965**	.958**
		Sig. (2-tailed)	.742	.129	.000	.	.000	.000
		N	14	14	14	14	14	14
	Ex of Good Lit?	Correlation Coefficient	.080	.405	.947**	.965**	1.000	.865**
		Sig. (2-tailed)	.787	.151	.000	.000	.	.000
		N	14	14	14	14	14	14
	How Literary Is This Poem?	Correlation Coefficient	.241	.513	.776**	.958**	.865**	1.000
		Sig. (2-tailed)	.407	.061	.001	.000	.000	.
		N	14	14	14	14	14	14

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 103.

Poem E: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.207	-.068	-.003	-.031	-.051
		Sig. (2-tailed)	.	.498	.825	.993	.920	.869
		N	13	13	13	13	13	13
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.207	1.000	-.068	.024	-.013	.042
		Sig. (2-tailed)	.498	.	.826	.938	.967	.893
		N	13	13	13	13	13	13
	Dixon Scale	Correlation Coefficient	-.068	-.068	1.000	.965**	.978**	.918**
		Sig. (2-tailed)	.825	.826	.	.000	.000	.000
		N	13	13	13	13	13	13
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.003	.024	.965**	1.000	.980**	.976**
		Sig. (2-tailed)	.993	.938	.000	.	.000	.000
		N	13	13	13	13	13	13
	Ex of Good Lit?	Correlation Coefficient	-.031	-.013	.978**	.980**	1.000	.936**
		Sig. (2-tailed)	.920	.967	.000	.000	.	.000
		N	13	13	13	13	13	13
	How Literary Is This Poem?	Correlation Coefficient	-.051	.042	.918**	.976**	.936**	1.000
		Sig. (2-tailed)	.869	.893	.000	.000	.000	.
		N	13	13	13	13	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 104.

Poem E: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.734**	-.304	-.389	-.269	-.469
		Sig. (2-tailed)	.	.001	.236	.123	.296	.058
		N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.734**	1.000	-.278	-.421	-.325	-.468
		Sig. (2-tailed)	.001	.	.280	.093	.203	.058
		N	17	17	17	17	17	17
	Dixon Scale	Correlation Coefficient	-.304	-.278	1.000	.873**	.901**	.686**
		Sig. (2-tailed)	.236	.280	.	.000	.000	.002
		N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.389	-.421	.873**	1.000	.951**	.886**
		Sig. (2-tailed)	.123	.093	.000	.	.000	.000
		N	17	17	17	17	17	17
	Ex of Good Lit?	Correlation Coefficient	-.269	-.325	.901**	.951**	1.000	.708**
		Sig. (2-tailed)	.296	.203	.000	.000	.	.001
		N	17	17	17	17	17	17
	How Literary Is This Poem?	Correlation Coefficient	-.469	-.468	.686**	.886**	.708**	1.000
		Sig. (2-tailed)	.058	.058	.002	.000	.001	.
		N	17	17	17	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 105.

Poem F: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.470	.056	-.099	-.102	-.283
		Sig. (2-tailed)	.	.105	.856	.747	.741	.350
		N	13	13	13	13	13	13
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.470	1.000	.554*	-.100	.024	-.287
		Sig. (2-tailed)	.105	.	.050	.744	.937	.342
		N	13	13	13	13	13	13
	Dixon Scale	Correlation Coefficient	.056	.554*	1.000	.494	.696**	.255
		Sig. (2-tailed)	.856	.050	.	.086	.008	.401
		N	13	13	13	13	13	13
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.099	-.100	.494	1.000	.932**	.866**
		Sig. (2-tailed)	.747	.744	.086	.	.000	.000
		N	13	13	13	13	13	13
	Ex of Good Lit?	Correlation Coefficient	-.102	.024	.696**	.932**	1.000	.727**
		Sig. (2-tailed)	.741	.937	.008	.000	.	.005
		N	13	13	13	13	13	13
	How Literary Is This Poem?	Correlation Coefficient	-.283	-.287	.255	.866**	.727**	1.000
		Sig. (2-tailed)	.350	.342	.401	.000	.005	.
		N	13	13	13	13	13	13

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

Fig. 106.

Poem F: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.839**	-.500*	-.585*	-.480	-.685**
		Sig. (2-tailed)	.	.000	.041	.014	.051	.002
		N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.839**	1.000	-.350	-.414	-.360	-.473
		Sig. (2-tailed)	.000	.	.168	.099	.155	.055
		N	17	17	17	17	17	17
	Dixon Scale	Correlation Coefficient	-.500*	-.350	1.000	.930**	.962**	.829**
		Sig. (2-tailed)	.041	.168	.	.000	.000	.000
		N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.585*	-.414	.930**	1.000	.971**	.956**
		Sig. (2-tailed)	.014	.099	.000	.	.000	.000
		N	17	17	17	17	17	17
	Ex of Good Lit?	Correlation Coefficient	-.480	-.360	.962**	.971**	1.000	.864**
		Sig. (2-tailed)	.051	.155	.000	.000	.	.000
		N	17	17	17	17	17	17
	How Literary Is This Poem?	Correlation Coefficient	-.685**	-.473	.829**	.956**	.864**	1.000
		Sig. (2-tailed)	.002	.055	.000	.000	.000	.
		N	17	17	17	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

<p style="text-align:center">Fig. 107.</p>

Poem G: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.683*	-.227	-.033	-.134	.156
		Sig. (2-tailed)	.	.010	.457	.916	.664	.611
		N	13	13	13	13	13	13
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.683*	1.000	-.118	.212	.077	.346
		Sig. (2-tailed)	.010	.	.701	.487	.802	.247
		N	13	13	13	13	13	13
	Dixon Scale	Correlation Coefficient	-.227	-.118	1.000	.606*	.800**	.450
		Sig. (2-tailed)	.457	.701	.	.028	.001	.123
		N	13	13	13	13	13	13
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.033	.212	.606*	1.000	.882**	.936**
		Sig. (2-tailed)	.916	.487	.028	.	.000	.000
		N	13	13	13	13	13	13
	Ex of Good Lit?	Correlation Coefficient	-.134	.077	.800**	.882**	1.000	.701**
		Sig. (2-tailed)	.664	.802	.001	.000	.	.008
		N	13	13	13	13	13	13
	How Literary Is This Poem?	Correlation Coefficient	.156	.346	.450	.936**	.701**	1.000
		Sig. (2-tailed)	.611	.247	.123	.000	.008	.
		N	13	13	13	13	13	13

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

<p style="text-align:center">Fig. 108.</p>

Poem G: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.671**	-.269	-.191	-.158	-.228
		Sig. (2-tailed)	.	.002	.280	.448	.531	.364
		N	18	18	18	18	18	18
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.671**	1.000	-.150	-.120	-.109	-.073
		Sig. (2-tailed)	.002	.	.551	.636	.668	.775
		N	18	18	18	18	18	18
	Dixon Scale	Correlation Coefficient	-.269	-.150	1.000	.886**	.929**	.746**
		Sig. (2-tailed)	.280	.551	.	.000	.000	.000
		N	18	18	18	18	18	18
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.191	-.120	.886**	1.000	.951**	.945**
		Sig. (2-tailed)	.448	.636	.000	.	.000	.000
		N	18	18	18	18	18	18
	Ex of Good Lit?	Correlation Coefficient	-.158	-.109	.929**	.951**	1.000	.816**
		Sig. (2-tailed)	.531	.668	.000	.000	.	.000
		N	18	18	18	18	18	18
	How Literary Is This Poem?	Correlation Coefficient	-.228	-.073	.746**	.945**	.816**	1.000
		Sig. (2-tailed)	.364	.775	.000	.000	.000	.
		N	18	18	18	18	18	18

**. Correlation is significant at the 0.01 level (2-tailed).

<p style="text-align:center">329</p>

Fig. 109.

Poem H: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.765**	-.544	.098	-.099	.182
		Sig. (2-tailed)	.	.002	.055	.750	.747	.552
		N	13	13	13	13	13	13
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.765**	1.000	-.142	-.054	-.099	-.216
		Sig. (2-tailed)	.002	.	.643	.862	.749	.479
		N	13	13	13	13	13	13
	Dixon Scale	Correlation Coefficient	-.544	-.142	1.000	.106	.218	-.135
		Sig. (2-tailed)	.055	.643	.	.731	.474	.659
		N	13	13	13	13	13	13
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	.098	-.054	.106	1.000	.921**	.851**
		Sig. (2-tailed)	.750	.862	.731	.	.000	.000
		N	13	13	13	13	13	13
	Ex of Good Lit?	Correlation Coefficient	-.099	-.099	.218	.921**	1.000	.633*
		Sig. (2-tailed)	.747	.749	.474	.000	.	.020
		N	13	13	13	13	13	13
	How Literary Is This Poem?	Correlation Coefficient	.182	-.216	-.135	.851**	.633*	1.000
		Sig. (2-tailed)	.552	.479	.659	.000	.020	.
		N	13	13	13	13	13	13

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Fig. 110.

Poem H: Spearman Rank Correlations--Number of Words Underlined vs. Literariness Scores (Non-Student Respondents)

			No. of Words Underlined (Mean Ranks)	No. of Underlinings (Mean Ranks)	Dixon Scale	Ex. of Good Lit? + How Literary?	Ex of Good Lit?	How Literary Is This Poem?
Spearman's rho	No. of Words Underlined (Mean Ranks)	Correlation Coefficient	1.000	.766**	-.440	-.396	-.422	-.438
		Sig. (2-tailed)	.	.000	.077	.116	.091	.079
		N	17	17	17	17	17	17
	No. of Underlinings (Mean Ranks)	Correlation Coefficient	.766**	1.000	-.508*	-.487*	-.487*	-.520*
		Sig. (2-tailed)	.000	.	.037	.047	.048	.033
		N	17	17	17	17	17	17
	Dixon Scale	Correlation Coefficient	-.440	-.508*	1.000	.878**	.951**	.798**
		Sig. (2-tailed)	.077	.037	.	.000	.000	.000
		N	17	17	17	17	17	17
	Ex. of Good Lit? + How Literary?	Correlation Coefficient	-.396	-.487*	.878**	1.000	.966**	.966**
		Sig. (2-tailed)	.116	.047	.000	.	.000	.000
		N	17	17	17	17	17	17
	Ex of Good Lit?	Correlation Coefficient	-.422	-.487*	.951**	.966**	1.000	.896**
		Sig. (2-tailed)	.091	.048	.000	.000	.	.000
		N	17	17	17	17	17	17
	How Literary Is This Poem?	Correlation Coefficient	-.438	-.520*	.798**	.966**	.896**	1.000
		Sig. (2-tailed)	.079	.033	.000	.000	.000	.
		N	17	17	17	17	17	17

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

For non-student respondents of Poem B, significant *negative* correlations were found

330

between the number of words underlined and impressions of literariness. The strength of these correlations was moderate-to-strong. We find another *negative* correlation for non-student respondents of Poem C, this time between the number of words underlined and the question "Is this poem an example of good literature?" Next, where Poem D is concerned, it is the students this time that produce *negative* correlations between number of words underlined and literariness scores. Next up is Poem F: for this poem, student data produced a moderate *positive* correlation between number of underlinings and their responses on the Dixon scale; however, this scale contains two items not specifically related to literariness. Meanwhile, the non-student data for this poem yielded significant *negative* correlations between number of words underlined and literariness scores. Finally, for Poem H, non-student data yielded significant *negative* correlations between the number of underlinings and literariness scores. In the above cases, the more respondents underlined, the less literary they judged the text to be.

So: did the poems score low on the literariness scale *in spite of* the items identified as striking? If so, perhaps there *were* striking words or phrases, but they did not pull their weight when it came time to evaluate the text as a whole. If this is the case, then perhaps some other quality intervened or overrode the experience of strikingness when it came to the literariness scales. Or did the items identified as striking actually have a *negative effect* on impressions of literariness? If we conceive of defamiliarisation and FG—as the early Russian Formalists and some of their successors arguably did—as "a category indicating 'essentials' of literariness in an absolute or material sense" (Van Peer 1986: 185),[57] this result would suggest a model in which literariness can be self-erasing. This is a strange inversion of a conundrum discussed earlier in connection with Russian Formalism: the problem of literature that no longer performs the function that defined it as literature to begin with. The difference is that in our

---

[57] To be clear, I am not suggesting that we should.

case, readers *are* attesting effects that would likely have been considered literary by some of the Russian Formalists, while at the same time these same effects are negatively correlated with the very quality that they—the effects—are supposed to create!

The fact that it is mostly non-students responding this way suggests a particular explanation. With few exceptions, older respondents (non-students) rated the texts lower in terms of impression of literariness than younger respondents (students). If the literariness questions worked the way I hoped, i.e. prompted readers to compare the text in front of them with other literary texts they had encountered in the past, then we might interpret the results as follows: whatever the basis for comparison was (a memory of some text they considered to be quintessentially literary, some constellation of feelings called up by the word 'literature' or 'literary,' or perhaps an image of a *kind* of text they thought might be considered literary by an expert), for the older respondents there was less of a 'match' between the *texts in front of them* and the *texts behind them*, so to speak. This, in turn, might suggest that younger respondents' basis for comparison was either more informed by or *more amenable to being informed by* Japanese literature. In other words, there was more of a match between the *texts in front of them* and the *texts behind them*, hence the higher literariness ratings—or they were more willing to adjust their conception of literature such that the *texts in front of them* fit into this category reasonably well. This is why I suggested that younger respondents' basis for comparison might be 'amenable to being informed by Japanese literature': it seems possible that there might exist in respondents' answers on the Likert scale what Austin (1975) called a "performative" element.

In contrast to constative utterances, performative utterances "do not describe but perform the action they designate," as Culler (1997: 95) explains. However, Austin ran into some trouble in maintaining a strict distinction between the two; as Culler (ibid.: 95-96) writes, "[t]he sentence 'The cat is on the mat,' your basic constative utterance, can be seen as the

elliptical version of 'I hereby affirm that the cat is on the mat,' a performative utterance that accomplishes the act of affirming to which it refers." In this way, the "contrast between constative and performative" is "redefined" (ibid.: 100). In other words, "once you allow for the existence of such 'implicit performatives,' where there is no explicitly performative verb, you have to admit that *any* utterance can be an implicit performative" (ibid.: 95). A statement like "[t]his poem is an example of good literature" may at a certain level look constative and therefore appear to "represent things as they are" (ibid.) in the speaker's world; but at another level the statement is also "imposing" these "linguistic categories" and thus may also be *re*-organising the world "rather than simply representing what it is" (ibid.: 100-101). In other words, respondents who strongly agree with the statement "[t]his poem is an example of good literature" might also be signalling that they think the poem *should* be considered an example of good literature. Thus, it is possible that the corresponding Likert items are also picking up, to a degree, what respondents think the category of literary texts *should be like*. For this reason, greater attention to the performative possibility of Likert items might result in finer-grained data in future studies—and that if we want a better picture of the kind of texts that respondents *have considered literary in the past*, it may do to rephrase certain questions to minimise this kind of ambiguity. For instance, one might include something like the following as a Likert item: "This poem resembles in important ways other texts that have been taught or presented to me as literature in the past."

In any case, the situation regarding student and non-student respondents I have described above is perhaps not very surprising—if we were to guess which age group would be most familiar with Japan and Japanese cultural products, we would probably point to the 18-to-24-year-old group in this survey. This is *not* to say that the results suggest that younger respondents' interpretations of the terms 'literary' or 'literature' are more plastic or adaptable than older respondents', or that they have stronger opinions about what should be deemed

literary. Indeed, the performative element may be equally present in the older group's responses: it is possible that the Likert-scaled items also picked up what older respondents' thought the category of literary texts *should be like*. Nevertheless, it would be interesting to know what specific reasons these two groups might have had for imposing these linguistic categories ("good literature," "literary") on the stimulus texts to the degree that they did. One possible explanation is that the two groups tended to value strikingness differently. If we recall back to H_4, Gordon, Lowder and Hoedemaker (2016: 165) found that "older adults perform as well or better than younger adults when higher-level meanings of a text are assessed." It might be that my request for respondents to underline striking words or phrases frustrated what was already a slower method of reading (associated with age-related cognitive changes) and tended to distract from the inclination to construct higher-level meanings from the texts, thereby resulting in lower scores for impression of literariness. In other words, the nature of the survey possibly slowed reading down a bit too much for some readers, frustrating expectations about what the reading process needs to be like in order for them to deem a text 'literary.'

Where the negative correlations between strikingness and literariness are concerned, it is tempting to posit a typology of literariness. In other words, perhaps the negative correlations above could be explained by invoking *two different kinds of literariness*: it could be that the literariness indexed by strikingness exists at the micro-level of textual perception, while the literariness indexed by the Likert-scaled items is an impression that emerges at the macro-level of textual evaluation. While one would normally expect the two to be highly correlated—making it difficult to spot the existence of *two different kinds of literariness* in the first place—perhaps for Poems B, C, D, F and H some unknown factor has nevertheless caused them to diverge for some readers. However, the results for the other poems show no positive correlations between strikingness and literariness. Moreover, while complexifying

the model by positing *two kinds of literariness* might be intellectually showy, it is not very parsimonious.

In any case, the results obtained here suggest that additional scales might be useful in future research. One might ask future respondents to indicate *why* they underlined particular words or phrases, and whether they felt such phrases to be typical of the text-type in question. One might also ask the degree to which such phrases contributed to their impression of the text as a whole. In this way, we might better understand which norms are more relevant to certain readers at certain points in the text, and how this background knowledge is activated.

H_6 concerns whether strikingness and enjoyment are correlated. H_a predicted that the *amount of FG* in a text would be positively correlated with enjoyment ratings. H_0 held that *amount of FG* would have no correlation with enjoyment ratings. The independent variable *amount of FG* was measured in two ways: first, by the number of FG features identified by my stylistic analyses of each text; and second, by the amount of *perceived FG*, as indicated by the number of words underlined for strikingness. The dependent variable *enjoyment* was measured by summing responses to two Likert-scaled items: "Did you enjoy reading this poem?" and "How enjoyable or aesthetically appealing did you find this poem?" I also took the opportunity to check whether literariness and enjoyment ratings were correlated, and to perform a principal components analysis (PCA) of the Dixon et al. (1993) appreciation scale, for reasons I discuss below.

The results for Surveys 1.1 and 1.2 show:

- **No significant correlation was found between *amount of FG* (as previously identified by stylistics analysis) and median enjoyment scores** (according to a Spearman rank test)
  - *this is evidence **against** the theory of FG*

- **No significant correlation was found between *amount of FG* (as measured in median number of words underlined per poem) and median enjoyment scores** (according to a Spearman rank test)
  - *this is evidence **against** the theory of FG*

- **Scores for enjoyment and literariness were significantly positively correlated for non-students for all of the poems** (according to a Spearman rank test)
  - *this is evidence **in favour of** the theory of FG*

- **Scores for enjoyment and literariness were significantly positively correlated for**

336

students for Poems C, D, E, G (according to a Spearman rank test)

- *this is evidence **in favour of** the theory of FG*

- **For students, there was no significant correlation between enjoyment and literariness scores for Poems A, B, H and F** (according to a Spearman rank test)

  - *this is evidence **against** the theory of FG*

Overall, we have found mixed evidence for this hypothesis. Some results support the theory of FG, and others do not. The motivation for including this hypothesis was the lack of strong evidence linking the perception of FG to positive aesthetic evaluations of literary texts. As I discussed earlier, Zyngier et al. (2007) demonstrated that the higher the overall amount of FG in a text (as quantified by stylistic analysis), the more likely readers were to perceive that text as complex. But they also hypothesised that complexity of patterning and depth of appreciation (as measured by the Dixon et al. [1993] appreciation scale) would be positively correlated. However, it was only in one of their three respondent groups that they found strong support for this claim. For reasons explained earlier in this thesis, I have not employed a re-reading protocol like the above two studies. Note also: I did not use the Dixon et al. (1993) appreciation scale to measure enjoyment, but rather summed the scores from a different combination of Likert items. This was because the Dixon scale combined items related to both literariness and enjoyment, and I was slightly sceptical about this construct of appreciation: it seemed possible to me that respondents could perceive a text as literary (according to their own understanding of that term), while not enjoying it very much at all, and vice versa. Thus, it seemed prudent to allow the reading subject to be a kind of divided subject here, capable of recognising where institutional values (how they *should* respond to texts) conflicted with their own values (how they *do* respond to texts). While the Dixon scale acknowledges that appreciation may be multifaceted, in combining items related to both literariness and enjoyment, their appreciation scale sums *across* this potential rift so that, for

337

instance, all other things being equal, a reader who rated a text 7 on literariness and 1 on enjoyment would produce the same score as a reader who rated a text 1 on literariness and 7 on enjoyment. The summed scores of these hypothetical readers would be equal, but would we then say that they had responded to that text the same way?

Where the split between institutional and individual modes of response is concerned, we might expect the strongest positive correlation between literariness and enjoyment to be found in the student group. As I touched on in my earlier discussion of the Horatian platitude ('instruct and delight'), the notion of enjoyment has long enjoyed pride of place in institutional constructions of the literary—at least in theory, if not in pedagogical practice. Thus, if students were the respondent group most immersed in the institution associated with this idea of literature, one might well expect their ratings for literariness and enjoyment to show a consistent relationship with one another. However, as the results below show, it is the *non-student group whose ratings for literariness and enjoyment are most highly correlated*.

One more consideration: the nature of this enjoyment. Some insights about this come from an unlikely source: the scholarship around horror spectatorship. One genre in which 'enjoyment' has attracted much theoretical elaboration is horror. It is not difficult to see that the so-called "paradox" of horror (cf. Strohl 2012: 211n1), i.e. readers/viewers finding hedonic value in negatively-valenced emotions, might also be implicated in aesthetic evaluations of texts from other genres—for instance, poetry that deals with unpleasant or taboo topics. The propositional content of Yoshihara's text comes to mind. After all, the speaker in this poem seems to suggest that all of humanity is ugly and should die out of the world. And yet some respondents rated it quite highly for enjoyment—in either translation. This poem also attracted one of the most intense comments written in the margin: one respondent indicated the first six stanzas of Miller and Kudo's translation and wrote "Beautiful," then drew a box around the last stanza—in which the speaker wields her fatal

imperative—and wrote "This ruins the poem." Nevertheless, they scored it a 6 on "Did you enjoy reading this poem?" and a 5 on "How enjoyable or aesthetically appealing did you find this poem?" for a total of 11 out of 14 possible points.

Strohl's (ibid.: 204) proposed solution to this problem of an ostensible mismatch between unpleasant content and positive reception is what he calls an "Aristotelian account of strong ambivalence […]." This concept refers to "taking pleasure in an experience partly *in virtue of* its painful aspects" (ibid.: 203; emphasis added). The problem with previous accounts, Strohl argues, is that they either deny that negative emotions make a positive contribution to the overall aesthetic experience, being something akin to emotional "collateral damage" (ibid.: 205), or they attempt to maintain that negative emotions are not really negative after all. Thus, it seems we need a theory that acknowledges that the negative feelings one experiences when engaging with such material *really might be negative or unpleasant or painful*, but that these perceptions nevertheless ultimately make some positive contribution to the overall experience of enjoyment. In Strohl's theory, pleasure concerns a "fit" between the "capacity" of a subject and the ability of an object to "fully activate" that capacity (ibid.: 208). This, for Aristotle, is what defined pleasure: it was a kind of relational effect in which two things were brought together in a complementary way. An example: "If someone is overheated […] water that would normally be unpleasantly cold may be pleasant to swim in" (ibid.).

While this account is more sophisticated than previous ones in acknowledging the complexity of the internal structure of aesthetic experiences, if crudely interpreted it would seem to suggest that audiences need to be in some initially unfulfilled capacity to obtain the eventual pleasurable outcome. Such an interpretation slides back all too easily into the kind of 'release valve' idea that audiences seek out such artwork to relieve themselves of unpleasant feelings. However, as Strohl (ibid.: 211n4) writes, "[c]ases where one enjoys the relief of a painful feeling are difficult to classify" with regard to horror spectatorship. It

would be one thing if "the pleasure horror audiences experience in connection with fear and disgust could plausibly be taken exhaustively to consist in the relief of these feelings, but this does not seem to be the case" (ibid.). After all, audiences "often report enjoying a horror film the most in the moments when it is most terrifying" (ibid.). Also, many horror films "offer no relief in the course of their narratives" (ibid.). I do not mean to push the comparison with Yoshihara's poem too far, but presumably it was at *the very end of the poem* where the above-mentioned respondent ("This ruins the poem") encountered the strongest unpleasant feeling. Like some texts in the horror or gothic tradition, Yoshihara's poem does not have a happy ending, at least in terms of propositional content. However, if the ending *really* ruined the affective experience of the poem, would this respondent have given it 11 out of 14 points for enjoyment? (The respondent was from the non-student group. As we can see from the tables below, the mean and median scores for Poem D for this group were 7.79 and 7.50 respectively.) Thus, as predicted by Strohl's theory, this respondent seemed to enjoy the Yoshihara poem *in spite of* there being no relief from the constant mentions of death that lead up to the final stanza. Thus, if this *lack of relief* was implicated in any of the scores, it seems more likely to have influenced impression of literariness. Indeed, this respondent gave the poem an 8 out of 14 on this quality (they recorded a 4 for both "In your opinion, how literary is this poem?" and "Is this poem an example of good literature?").

Fig. 111. Enjoyment Scores (Students)

	Poem A	Poem B	Poem C	Poem D	Poem E	Poem F	Poem G	Poem H
Mean Enjoyment Score	8.59	10.65	9.12	10.44	9.92	10.54	9.31	10.00
Median Enjoyment Score	9.00	11.00	10.00	10.50	9.00	11.00	9.00	11.00
SD (Sample)	2.81	1.54	3.26	1.75	2.47	2.76	2.10	3.00
SD (Population)	2.72	1.49	3.16	1.69	2.37	2.65	2.01	2.88

Fig. 112. Enjoyment Scores (Non-Students)

	Poem A	Poem B	Poem C	Poem D	Poem E	Poem F	Poem G	Poem H
Mean Enjoyment Score	6.40	8.92	8.62	7.79	10.06	9.35	10.78	6.65
Median Enjoyment Score	6.00	9.50	9.00	7.50	11.00	10.00	12.00	7.00
SD (Sample)	3.60	4.06	3.36	3.70	3.05	3.35	2.18	3.22
SD (Population)	3.48	3.88	3.22	3.57	2.96	3.25	2.12	3.12

The above respondent was somewhat atypical for the non-student group: this respondent's enjoyment score for Poem D was about a full standard deviation away from the mean. Meanwhile, they indicated an enjoyment score of 12 out of 14 points for Poem A, which is more than a full standard deviation from the mean. One insight produced by the tables above is: non-students attested lower enjoyment scores than students for poems that explicitly mention death or dying. The texts that explicitly mentioned death or dying were Poems A ("dead of night"), D (throughout), and H ("And death is what we need"). The theory of FG thus far has not had much to say about enjoyment per se, so it is unclear whether this result should be taken as evidence *for* or *against* the theory. But it seems pertinent that all student respondents were in the youngest age category, whereas the non-student group included many respondents in higher age categories.

To determine whether the numerical differences are statistically significant, I ran a Mann-Whitney U test. If we treat this as a kind of exploratory study (to my knowledge no other study has tested whether younger and older respondents differ in terms of textual preference along this variable), then it is permissible to set a slightly less restrictive threshold for significance (Mazurek Melnyk, Morrison-Beedy and Moore 2012: 55) such that $p = 0.10$. In fact, all results were significant, which means that older respondents tended to report lower enjoyment scores for poems that mentioned death (with the results for Poem H being especially statistically significant, such that $p = .004$).

It is not possible to say for certain that the lexical content related to death is responsible

for this result; the study was not designed with this in mind. But the results suggest that this may be an area of potential interest to those researching in the intersection of library science and gerontology. As Harvey and Dutton wrote in 1979, "There is a paucity of research concerning the reading interests of older adults" (Harvey and Dutton 1979: 209). Unfortunately, despite the efforts of some scholars, the situation does not seem to have changed much since then. Nevertheless, it is not difficult to see how knowledge about the reading habits and preferences of older adults might "be helpful in informing the general public about the richness of the intellectual activities of older persons" (Smith 1993: 426). Moreover, this knowledge might also be "useful for library services planning, curriculum development for educational programs aimed at the older or lifelong learner, and a variety of social and community programs whose collective goal is to enhance the lives of aging adults" (ibid.).

Overall, where non-student respondents were concerned, the two most highly rated texts for enjoyment were Poems E and G. It is interesting that Poem E should be rated so highly, as it and Poem A are translations of the same source text (Ishigaki's "Shijimi"), but Poem A was among the least enjoyed poems. Could it be because—compared to Poem A, which includes the words "dead," "cackle" and "evil"—Poem E includes fewer negatively-valenced words? *If so, it might be the case that emotional valence is a better predictor of enjoyment for older readers than younger ones, as there is much less of a difference in enjoyment scores between Poems A and E for student respondents*. This might also explain why older respondents preferred Poem G over Poem C, both translations of a tanka by Yosano Akiko. While both texts included positively-valenced vocabulary, Poem G arguably has *more* positively-valenced vocabulary (Poem C has "merciful" and "love," while Poem G has "love's," "desire" and "Dear"). It may also have helped that Poem G has a more typical layout than Poem C. As one non-student from the 65+ category wrote about Poem G in Survey 2.2:

"classic west european layout. could have been written by john donne! [sic]" In any case, the above discussion suggests some qualitative differences between the student and non-student groups in terms of enjoyment—it is wise not to generalise too much from these results, but they have the benefit of suggesting additional falsifiable hypotheses for further research in a socially consequential research area.

Next, I discuss the results of my Spearman rank tests. First, I wanted to determine whether *amount of FG* as previously identified by my stylistics analysis was correlated at all with median enjoyment scores. To do this, I produced two rankings of the poems based on these two measures and used a Spearman rank test to compare the rankings. No significant correlation was found. It may indeed be true that texts with more FG are perceived as more complex by respondents; however, the claim that respondents are more likely to rate complex texts highly is not borne out here, at least in my survey protocol. Note: all respondents were analysed as one group for this test. Also: because Poems D and H did not undergo stylistic analysis, they had to be excluded from this test. Nevertheless, since this result goes against the theory of FG, I have highlighted it in yellow.

Fig. 113. Spearman Test: FG Features (Stylistic Analysis) vs. Enjoyment Scores

			Mean Rank by FG Features	Mean Rank by Median Enjoyment Score
Spearman's rho	Mean Rank by FG Features	Correlation Coefficient	1.000	.273
		Sig. (2-tailed)	.	.600
		N	6	6
	Mean Rank by Median Enjoyment Score	Correlation Coefficient	.273	1.000
		Sig. (2-tailed)	.600	.
		N	6	6

Next, I ran another Spearman rank test to see whether the *amount of FG* (as measured in median number of words underlined per poem) was correlated in any way with enjoyment scores. Once again, I analysed all respondents as one group. This time, however, I was able to include all the stimulus texts. As the results below show, no significant correlation was found. Since this result does not support the theory of FG, I have highlighted it in <mark>yellow</mark>.

Fig. 114. Spearman Test: No. Words Underlined vs. Enjoyment Scores

			Mean Rank by Number of Words Underlined	Mean Rank by Median Enjoyment Score
Spearman's rho	Mean Rank by Number of Words Underlined	Correlation Coefficient	1.000	.530
		Sig. (2-tailed)	.	.177
		N	8	8
	Mean Rank by Median Enjoyment Score	Correlation Coefficient	.530	1.000
		Sig. (2-tailed)	.177	.
		N	8	8

Finally, I performed Spearman rank tests to determine whether there were any correlations between ratings for enjoyment and ratings for impression of literariness. To do this, I combined the two items related to enjoyment, and the two relating to literariness. As we can see from the results below, for the non-student group, enjoyment and impression of literariness were strongly correlated for every poem—sometimes *very* strongly correlated, as in Poems A and B. The more non-students enjoyed a poem, the more likely they were to rate it highly for impression of literariness. Likewise, the less they enjoyed a poem, the more likely they were to give it a low literariness score. However, where the student group was concerned, such correlations only existed in half the texts, and these correlations on the whole were not as strong. This suggests a difference in the way that students and non-students

344

responded to the poems: **the two qualities *enjoyment* and *impression of literariness* were somewhat less tied together for student respondents, whereas these two qualities seemed to move in unison for the non-student group.** However, the direction of causation is difficult to determine: did the non-students consider texts literary *because* they enjoyed them to the extent that they did? Or did they enjoy the texts to the extent that they did *because* they deemed them literary?

Fig. 115. Spearman Rank Correlations—Enjoyment vs. Literariness (Students)

Poem	Correlation coefficient	p-value	N
A	0.404	0.108	17
B	0.319	0.212	17
C	0.654	0.004	17
D	0.693	0.003	16
E	0.836	0.000	13
F	0.268	0.377	13
G	0.627	0.022	13
H	-0.178	0.560	13

Fig. 116. Spearman Rank Correlations—Enjoyment vs. Literariness (Non-Students)

Poem	Correlation coefficient	p-value	N
A	0.938	0.000	15
B	0.952	0.000	12
C	0.628	0.021	13
D	0.893	0.000	14
E	0.836	0.000	17
F	0.853	0.000	17
G	0.771	0.000	18
H	0.883	0.000	17

I also took the opportunity to perform a PCA of Dixon et al.'s (1993) appreciation scale.

Studies using this scale have made an important, unspoken assumption: namely, that the variables represented by the questions are unidimensional, and that they therefore measure a single theoretical construct (in this case, 'appreciation'). To a certain way of thinking, the assumption of unidimensionality is not unreasonable; intuitively, it does not seem wrong that people would be likely to enjoy texts that they also deem 'good,' nor does it seem controversial that readers would recommend a text they enjoy. (But then, as one respondent cheekily wrote in the margin: "I don't usually go around recommending poems.") Nevertheless, it seemed prudent to investigate this matter further, supplementing the scale with an additional question related to *impression of literariness* and an additional question related to *enjoyment*. It seemed possible that for some readers, a particular stimulus text might sufficiently resemble other texts they had encountered before as literature, and yet not make for particularly enjoyable reading. I found promising evidence that this is the case from my PCA.

PCA seeks to reduce a larger set of variables to a smaller set, based on how correlated the original variables were. The first results I produced were for Survey 1.1. I performed separate PCAs for non-student and student responses. Both sets of data showed *unidimensionality among all of the questions*, whether they were related specifically to enjoyment or impression of literariness. **In other words,** *for Survey 1.1, all five Likert questions seemed to be measuring the same underlying variable.* Moreover, the correlations held for each text. It was not until I performed the PCA for student respondents to Survey 1.2 that a somewhat different picture emerged.

For Survey 1.2, the component matrix for the non-student group (below) identified only one component—with all the variables loading quite highly. Moreover, the total variance table for this group shows that this one component has high explanatory power, accounting for 88.47 percent of the variance in the survey responses. However, the table for the student

group shows that *two components* have been identified. In this table, three variables load highly onto Component 1: the two items dealing specifically with enjoyment and the one related to 'action tendency.' As the value of Component 1 goes up, these three variables tend to move in the same upward direction. Meanwhile, three variables also load well onto Component 2. Two of these concern impression of literariness. The third is an enjoyment question which also loads highly on Component 1. As the value of Component 2 goes up, these three variables move together in a downward direction. The fact that the question "Did you enjoy this poem?" loads on both components is problematic. This could indicate that it measures two different constructs at the same time—a situation which calls its usefulness as a variable into question.

Fig. 117. Component Matrix for Survey 1.2 (Non-Students)

Component Matrix^a

	Component 1
ExGoodLit	0.975
EnjReading	0.958
Aesthetic	0.952
Recommend	0.942
HowLit	0.874

Extraction Method: Principal Component Analysis.
a. 1 components extracted.

Fig. 118. Survey 1.2.: Total Variance—Non-Students

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.424	88.474	88.474	4.424	88.474	88.474
2	0.327	6.536	95.010			
3	0.117	2.334	97.344			
4	0.089	1.774	99.117			
5	0.044	0.883	100.000			

Extraction Method: Principal Component Analysis.

Fig. 119. Pattern Matrix for Survey 1.2 (Students)

Pattern Matrix[a]

	Component	
	1	2
Aesthetic	.890	.138
Recommend	.861	-.046
EnjReading	.719	-.317
HowLit	-.070	-.981
ExGoodLit	.109	-.925

Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization.[a]

a. Rotation converged in 6 iterations.

Fig. 120. Survey 1.2.: Total Variance—Students

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	3.007	60.136	60.136	3.007	60.136	60.136	2.478
2	1.134	22.684	82.820	1.134	22.684	82.820	2.370
3	.475	9.505	92.325				
4	.296	5.924	98.249				
5	.088	1.751	100.000				

Extraction Method: Principal Component Analysis.

a. When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

Thus, the PCA for the student group of Survey 1.2 suggests that *there are two distinct constructs being measured.* The first construct we can call 'enjoyment.' The second one is difficult to name, owing to the way the Pattern Matrix data are reported—as the value of this component goes up, scores for impression of literariness go down (while scores for enjoyment are relatively unaffected). However, the identification of this component reflects scores for Poems E, F, G and H in aggregate. Spearman tests showed that for poems F and H, there was no correlation between *enjoyment* and *impression of literariness*. For poem H, the loadings on Component 1 (see below) suggest that the enjoyment variables all move together. Meanwhile, the loadings on Component 2 suggest that as scores for *impression of literariness* go up, scores for appreciation are relatively unaffected.[58] Again, this result should be approached with some caution owing to the small sample size.

---

[58] The PCA for Poem F has a negative eigenvalue and thus SPSS halts the analysis.

Fig. 121. Pattern Matrix for Poem H (Students)

Pattern Matrix[a]

	Component	
	1	2
EnjReading	0.897	-0.006
Recommend	0.859	-0.226
Aesthetic	0.825	0.124
ExGoodLit	0.118	0.939
HowLit	-0.177	0.862

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
a. Rotation converged in 4 iterations.

These are interesting results which bear on future research directions. While we cannot make statistical inferences to any population other than the one sampled above (the student group for Survey 1.2), the results suggest that: under the right circumstances, the above questions measure two separate constructs (*enjoyment* and *impression of literariness*); moreover, these two constructs are not necessarily positively correlated. Indeed, as the table below for Poem H suggests, the two constructs may even be negatively correlated. As Component 1 (*enjoyment*) goes up, Component 2 (*impression of literariness*) goes down, and vice versa.

Fig. 122. Component Correlation: Poem H

Component Correlation Matrix

Component	1	2
1	1.000	-0.239
2	-0.239	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

As for why only the student group attested these results, we can propose a tentative hypothesis: there may be an awareness on the part of these students that what they enjoy reading and what they acknowledge as belonging to the category of literature may not always perfectly align. For these readers, enjoyment may not necessarily be a big part of literary reading: this might reflect their experiences of engaging with literature in a more academic way (i.e. having texts assigned to them for purposes other than appreciation). From the standpoint of students in a World Literature class (which these students were), this result might also indicate an awareness of their own cultural-historical situatedness as readers from a different literary milieu—an awareness, in other words, that their enjoyment of the TT may depend in part on familiarity with the source-language and culture(s), and familiarity with the ways in which the ST has been read and appreciated in that language. Thus, despite not necessarily enjoying a particular text, these readers may still be willing to recognise it as belonging to some category of literature. In terms of *intentio lectoris*, this pattern of response would seem to testify to a kind of divided subjectivity—one that is forthcoming about the degree to which it enjoyed (or did not enjoy) the TT, but willing to allow its perception of literariness to be informed by the reception(s) of the ST in the source language. This 'benefit of the doubt,' as it were, regarding literariness might then serve as a kind of impetus for the TT reader to consider the perspectives and affectivities of a ST reader, with the potential for increased enjoyment of the text. **Thus, we might conceive of reader response here as a system involving both texts.**

H_7 is the first of three hypotheses concerning Venuti's theory of foreignisation. This hypothesis concerned *impression of foreignness*, as measured by summing scores for the three additional Likert-scaled items in Survey 2. H_a predicted that texts appearing in the 'ostensibly foreignized' condition would elicit higher scores for impression of foreignness. H_0 stated that whether a text was classified as 'ostensibly foreignising' or 'ostensibly domesticating' would make no difference when it came to foreignness scores. If we treat this as a kind of exploratory study (to my knowledge no other study has operationalised concepts from the theory of foreignisation), then it is permissible to set a slightly less restrictive threshold for significance (Mazurek Melnyk, Morrison-Beedy and Moore 2012: 55) such that $p = 0.10$.

Results for Survey 2 show:

- **for every pair of 'competing' translations, the text classified as 'ostensibly foreignising' had a higher median foreignisation score** (according to descriptive statistics)

  - *this is evidence **in favour of** the theory of foreignisation*

- **for every pair of competing translations but one (Poems A and E), the 'ostensibly foreignising' one nevertheless has more variability in the foreignisation scores** (according to descriptive statistics)

  - *this is evidence **against** the theory of foreignisation*

- **only one of the 'ostensibly foreignising' texts (Poem C) scored significantly higher than its ostensibly domesticating counterpart (Poem G)** (according to a rank biserial correlation)

  - *this is evidence **against** the theory of foreignisation*

- **text condition was not systematically related to impression of foreignness** (according to Friedman two-way ANOVA)
  - *this is evidence **against** the theory of foreignisation*
- **no correlation was found between number of items underlined for impression of foreignness and overall foreignness scores** (according to a Spearman rank correlation)
  - *this suggests further survey refinements are necessary*

All in all, we have *not* found good evidence in favour of the theory of foreignisation for this hypothesis. First, let us discuss the positive result: for every pair of competing translations (A/E, B/F, C/G, D/H), the text that had been classified as 'ostensibly foreignising' beforehand yielded a higher median foreignisation score than the ostensibly domesticating text. Note: the median is robust to outliers, while the mean is not, and there was one outlier in the results for Poem G. As a kind of preliminary result, we could say that this is evidence in favour of the theory of foreignisation.

Fig. 123. Foreignisation Scores (All Respondents)

	Poem A	Poem B	Poem C	Poem D	Poem E	Poem F	Poem G	Poem H
Mean Foreignisation Score	12.50	12.58	12.42	12.33	12.67	12.33	9.78	13.22
Median Foreignisation Score	12.50	11.50	12.50	13.00	11.00	12.00	9.00	14.00
SD (Sample)	2.02	2.54	3.32	1.97	3.32	3.08	2.54	2.64
SD (Population)	1.94	2.43	3.17	1.89	3.13	2.91	2.39	2.48
Text Condition (For/Dom)	For	Dom	For	Dom	Dom	For	Dom	For

However, as we can see from the standard deviations, for every pair of competing translations (except Poems A/E), *it is the ostensibly foreignising one that has more variability in the foreignisation scores*. This implies that there was less agreement among respondents about the impression of foreignness of these texts. It is somewhat paradoxical: the above

results suggest respondents agreed more about *where a particular effect was not*, than about *where it was*. Nevertheless, as this result goes against the universalising assumptions of the theory, this is evidence against the theory. If it is on the basis of a text's effect on readers that we designate it as foreignising, then the effect in question turned out to be the most variable for the *very texts we would label as foreignising here* (based on one text in the pair having a higher median impression of foreignness score.) The issue seems unresolvable in Venuti's framework without either (1) claiming that some readers have responded the wrong way, i.e. criticising their *intentio lectoris*; or (2) shifting the basis of the designation away from effect toward some notion of *intentio operis* or *intentio translatoris*. However, if we recall that "it is possible to speak of text intention only as the result of a *conjecture* on the part of the reader" (Eco 1990: 58), then at the end of the day, this is not too different from claiming that the reader's conjecture about the text was wrong, thus, the reader has responded the wrong way. We have come full circle to de-emphasising *intentio lectoris* again, on the basis of an appeal to the reader's expertise (or lack thereof).

In any case, while the higher median impression of foreignness scores are a promising initial result, we need to determine whether the differences between the 'competing' poems are statistically significant. To do this, we may perform a rank biserial test to determine whether text condition is associated with impression of foreignness scores. A positive correlation value on this test indicates that the *higher* foreignisation score of the pair is associated with the 'ostensibly foreignising' text. However, only one of the results was significant at the level of $p < 0.10$. This is the result for Poems C and G (see table below). This results shows that only one of the 'ostensibly foreignising' texts (Poem C) scored significantly higher than its ostensibly domesticating counterpart. This is evidence *against* the theory of foreignisation.

Fig. 124. Rank Biserial Correlation: Poem C vs. Poem G

			Text Condition (Foreignising vs. Domesticating)	Foreignisation Score
Spearman's rho	Text Condition (Foreignising vs. Domesticating)	Correlation Coefficient	1.000	.432
		Sig. (2-tailed)	.	.051
		N	21	21
	Foreignisation Score	Correlation Coefficient	.432	1.000
		Sig. (2-tailed)	.051	.
		N	21	21

Note: in no survey did any respondent directly compare competing translations. Each rank biserial result thus represents data from two surveys, with one poem coming from Survey 2.1 (ABCD) and the other coming from Survey 2.2 (EFGH). Since the above test involved comparing scores *between* surveys, I also compared scores *within* each survey (i.e., compared Poems A, B, C and D with each other, and compared Poems E, F, G and H with each other). This was done with a Friedman two-way ANOVA. A significant result on this test would suggest that one or more texts were consistently rated higher than others for impression of foreignness. This would be a result in favour of the theory. However, the test for Survey 2.1 revealed that the mean ranks of each poem (ranked by impression of foreignness) are very close. Moreover, as the *p*-value for this test was higher than 0.10 ($p = 0.980$), we cannot reject the null hypothesis of equal mean ranks. This suggests that text condition does not have any relationship with overall impression of foreignness scores for Survey 2.1.

The situation regarding Survey 2.2 is slightly different: a look at the ranks suggests that

one of our poems (Poem G) has underperformed on impression of foreignness. As this was one of the ostensibly domesticating poems, this makes sense in terms of the theory of foreignisation. However, post-hoc testing reveals that this text underperforms in a statistically significant way *in relation to only one of the two ostensibly foreignising texts (Poem H).* Moreover, it underperforms to almost the same degree of statistical significance in relation to the *other ostensibly domesticating text (Poem E).* For this reason, I have resisted highlighting the associated *p*-value in green. Since no overall pattern of underperformance emerges for the other ostensibly domesticating poems, we must conclude that we do not have good evidence that text condition and impression of foreignness are associated.

Fig. 125. Friedman Test: Survey 2.2

Ranks

	Mean Rank
Impression of Foreignness: Poem E	2.89
Impression of Foreignness: Poem F	2.72
Impression of Foreignness: Poem G	1.44
Impression of Foreignness: Poem H	2.94

Test Statistics[a]

N	9
Chi-Square	8.547
df	3
Asymp. Sig.	.036

a. Friedman Test

If identifiable textual features contribute to the overall impression of foreignness of the text, then to one way of thinking, the texts with the highest foreignness scores should also be

the ones with the most items underlined. To test this, we can perform a Spearman rank correlation. First, we rank texts in terms of the median number of items underlined, then we make another list which ranks them by total foreignness score, then we compare the rankings. We can see these rankings below—they do not appear to match very well, apart from the bottom two poems in each table.

Fig. 126. Poems Ranked by Median Number Items Underlined

Poem	Median Num. Words Underlined	Mean Rank
Poem F	5.00	1
Poem B	4.00	2
Poem A	3.00	3
Poem C	2.00	4.5
Poem H	2.00	4.5
Poem D	1.00	6
Poem E	0.00	7.5
Poem G	0.00	7.5

Fig. 127. Poems Ranked by Median Foreignness Score

Poem	Median Foreign-ness Score	Mean Ran
Poem H	14.00	1
Poem D	13.00	2
Poem A	12.50	3.5
Poem C	12.50	3.5
Poem F	12.00	5
Poem B	11.50	6
Poem E	11.00	7
Poem G	9.00	8

A Spearman rank test confirms our suspicions: no correlation between the two sets of rankings was found (such that rho = 0.248 and p = 0.553). **In other words, the number of**

words underlined for impression of foreignness does not seem to have anything to do with the overall impression of foreignness score (the Likert questions). Two possibilities emerge: not all items underlined gave the same *degree* of impression of foreignness, and/or the overall impression of foreignness of a text goes beyond identifiable textual features (or at least the kind of features amenable to underlining.)

Support for the latter position is found in comments left by respondents. As one respondent indicated in the margin of Poem E, it was "not words" but "meter" that gave an impression of foreignness. Another respondent remarked of the same text that they "found the content of [Poem E] foreign but not its use of language." Thus, as was the case with strikingness, a simple additive model in which each underlined item 'weighs' the same in the final analysis is not enough to gain a full picture of the phenomenon. Thus, in future studies it may help to ask respondents to quantify *how much* of an impression of foreignness each underlined item gave, and *why* they underlined each item.

Finally, we may discuss results from a qualitative perspective. When we look at the words most frequently underlined in Survey 2 for impression of foreignness (see tables below), several intriguing phenomena stand out. Some of these corroborate Venuti's theory, some are difficult to account for within Venuti's theory, while still others suggest an theoretical overlap with strikingness.

Fig. 128. Poem A: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5
	1	shijimi	1	dead								
			1	in								
			1	night								
			1	of								
			1	the								
			1	up								
			1	woke								
Non-Students												
	2	shijimi	1	night	1	kitchen	1	clams				

Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10	Freq	Line 11
								3	had	3	half
								3	help	3	mouth
								3	it	3	open
								2	couldn't	2	all
								2	to	2	night
								1	after	2	sleep
								1	that	2	with
1	gobble					1	evil			1	half
1	up									1	mouth
1	you									1	night
										1	open

359

Fig. 129. Poem B: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7
	1	isolation	1	its			1	Mars	1	on	2	do	1	and	1	at
	1	milliard							1	their	2	don't	1	sloop	1	but
	1	of									2	i	1	wike	1	times
	1	the									2	know	1	wook		
											2	they				
											2	what				
Non-Students																
	2	milliard											2	sloop	1	Earth
													1	wike		
													1	wook		

Freq	Line 8	Freq	Line 9	Freq	Line 10	Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16
				8	of							1	all	1	milliard	1	involuntary
				7	the							1	are			1	sneeze
				4	force							1	so				
				4	isolation							1	uneasy				
				3	is							1	we				
				3	pulling												
				3	together												
		1	universal			1	universe	1	our	1	universe			2	milliard	1	light
								1	unite								
								1	wants								

Fig. 130. Poem C: Most Frequently Underlined Words (Foreignness)

Students	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5
	1	merciful	2	traveler			2	a	4	take
							2	ask	3	a
							2	don't	3	room
							2	for	2	but
							2	shortcut	2	me
									2	with
Non-Students										
	2	merciful	3	traveler					1	room
	1	is								
	1	rainfall								

Fig. 131. Poem D: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
			1	dies					1	dies			1	lifts								
			1	sea					1	moon												
			1	the					1	the												
Non-Students			Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
			1	sea					1	moon												

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
2	dying	1	again	1	beauty	1	dies			1	remains			1	undarkened					4	dead
2	repeatedly	1	comes	1	if	1	of													4	drop
		1	time	1	the	1	sea													4	people
						1	that														
						1	the														
Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21

Fig. 132. Poem E: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7
	1	Shellfish					2	alive	1	all	1	hag's				
							2	mouths	1	breaks	1	laugh				
							2	open	1	day						
							2	slightly	1	eat						
							2	their	1	them						
							2	with	1	when						
							1	are								
Non-Students																
					1	little					1	a				
					1	shellfish					1	hag's				
											1	laugh				

Fig. 133. Poem F: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9
			1	orb			2	Mars			1	Martians	2	hararaing	1	Earth			1	of
			1	small			1	for					2	kiruruing	1	friends			1	power
							1	friends					2	neririing	1	have			1	solitutdes
							1	on					2	or	1	like				
							1	wish							1	on				
															1	they				
															1	to				
Non-Students																				
			2	orb	2	waken					1	orb	4	hararaing					1	gravitation
			1	small									4	kiruruing					1	is
													4	neririing					1	of
													4	or					1	power
																			1	solitudes
																			1	the
																			1	universal

Freq	Line 10	Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16
				1	for								
				1	seek								

Fig. 134. Poem G: Most Frequently Underlined Words (Foreignness)

Students	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5
	1	rain	1	desire	N/A	N/A	N/A	N/A	1	here
			1	of					1	lodging
			1	love's					1	make
									1	your
Non-Students										
	1	rain	1	desire	1	traveler	N/A	N/A	N/A	N/A
			1	of						
			1	love's						

362

Fig. 135. Poem H: Most Frequently Underlined Words (Foreignness)

Students	Freq	Title	Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
	N/A	N/A	1	sea	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Non-Students			Freq	Line 1	Freq	Line 2	Freq	Line 3	Freq	Line 4	Freq	Line 5	Freq	Line 6	Freq	Line 7	Freq	Line 8	Freq	Line 9	Freq	Line 10
	1	瞬間	1	is	1	日が	1	暮れる	2	fatality	1	夜が	1	開ける	2	fatality	1	人も	1	死ね	N/A	N/A
			1	sea	1	a			1	is	1	day			1	a	1	dead				
			1	terminal	1	casualty			1	moon	1	is			1	is	1	goes				
			1	the	1	day			1	the	1	terminal			1	time						
					1	is					1	the										
					1	the																

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
1	always	1	generous	1	if	1	and	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	1	人も	1	死ね
1	back	1	is	1	majesty	1	of														
1	but	1	itself	1	the	1	moon														
1	coming	1	time			1	sea														
		1	with			1	the														

Freq	Line 11	Freq	Line 12	Freq	Line 13	Freq	Line 14	Freq	Line 15	Freq	Line 16	Freq	Line 17	Freq	Line 18	Freq	Line 19	Freq	Line 20	Freq	Line 21
1	時が	1	またくる	N/A	N/A	1	うつ	1	くしさ	N/A	N/A	1	のこる	1	なら	N/A	N/A	1	人も	1	死ね
								1	across											1	and
								1	saved											1	death
																				1	is
																				1	need
																				1	we
																				1	what

The first phenomenon to mention is perhaps the least expected—this is the low incidence of respondents underling actual Japanese vocabulary, either in the form of transliterated loans (like "Shijimi" in Poem A), or items preserved in their original Japanese script (as in Poem H). Overall, when presented with poems featuring both Japanese and English vocabulary (as in Poems A and H), participants tended to underline English-language items, when they underlined anything at all.

The likeliest explanation has to do with the nature of the population sampled: where student respondents were concerned, some respondents were taking a level-two Chinese literature class. So perhaps when it came to Poem H, the Japanese script used in that text (particularly the kanji) seemed sufficiently similar to Chinese script that it did not feel foreign

to them, hence they did not underline it. Although, if this were the case, it would fail to explain why no items in the hiragana syllabary were underlined: the characters in this syllabary are unique to Japanese. Script similarity between Chinese and Japanese would also not explain why only one student respondent underlined "Shijimi" in Poem A—that item was in English script. Of course, it is not out of the question that some of these students may have also studied Japanese. This would raise an interesting possibility—if we recall that only responses of native English speakers were accepted for Surveys 1 and 2, then *even if these respondents had studied Japanese*, these results would suggest that they felt the use of their native language was *more foreign* than the other language. It is a limitation of my survey that I did not include a question about which languages respondents had studied. If it could be shown that some respondents had studied Japanese, then the results would partially confirm Folkart's (2007: 299) position: "*Writing*, in other words, *is inherently foreign*. Literature— poetry—is *always written from just outside the idiom*." Perhaps this is what respondents' results show: that *this* is the kind of foreignness that really matters. Perhaps Folkart is right when she suggests that the relevant "esthetic [...] information" in a translation is "conveyed at levels far above the grain of the text" (ibid: 7)—the grain being where, she argues, the "focus" of foreignisation tends to fall (ibid.: xi). This would explain why so few respondents underlined either transliterated or non-transliterated Japanese for strikingness or impression of foreignness. The main exception is the Japanese-derived mock-Martian in Poem F. But perhaps readers simply felt these to be more aesthetically relevant than other Japanese words in the poems.

Then again, the survey instructions may have also played a part. I purposefully left it up to respondents to decide how to interpret 'foreign,' in the hopes of casting as wide a net as possible to study all the different ways this term was understood. This strategy paid off here, as the results suggested that there were four patterns by which respondents interpreted the

instructions to underline items that gave an impression of foreignness. It struck me that underlinings could be divided into four tentative categories: 1) nonstandard items, 2) phrases that might sound unnatural to native English speakers, 3) words or phrases that were striking to respondents (i.e. items also underlined in Survey 1) and 4) foreign vocabulary.

In the first category are items like 'traveler,' 'milliard,' 'wike,' 'wook.' 'Traveler' is the American spelling; one respondent even jotted 'sp' (spelling) next to their underlining. The *OED* remarks that 'milliard' is "now largely superseded by billion"; meanwhile, 'wike' is presented as a Martian word, though it also existed historically. Thus, 'traveler' is an example of synchronic variation, while 'milliard' is an example of diachronic variation. In the second category, we might put items like 'seek for' or 'the pulling together of the force of isolation.' In the third category would be items underlined in both Survey 1 and 2: 'the sea dies,' 'the sea is terminal,' etc. In the fourth category would be words like 'Shijimi' or the Japanese characters from Poem H. These categories are not exhaustive or even mutually exclusive, and other analyses are possible; however, categories 1, 2 and 4 seem to coincide (to a degree) with Venuti's descriptions of the foreign in Anglophone reception scenarios. Meanwhile, category 3 seems to bolster Folkart's position that viable poetic language already produces a sense of the foreign. Thus, it seems that while several respondents indicated difficulty understanding what was meant by 'foreignness,' they nevertheless interpreted it in one or more of the senses employed by Venuti. Given the above patterns, one task for future survey research is to frame tasks such that respondents do not feel limited to just one or two of these interpretations of 'foreignness.' Moreover, while I lack the space here to discuss responses from L2 English speakers, this might also be a fruitful avenue for research. (One such respondent underlined every word in Poem D and remarked to me later, "They're all foreign to me," apparently mobilising her L1 linguistic identity for this poem but, curiously, not the others.)

5.2.8. HYPOTHESIS 8 (FOREIGNISING FEATURES)

H$_8$ was the second hypothesis related to foreignisation. H$_a$ predicted that readers would primarily underline those features that can be described as foreignising according to the criteria discussed in section 1.3 and in Chapter 3. As I discussed at the end of the previous section, the data suggested four initial categories: 1) nonstandard items, 2) phrases that might sound unnatural to native English speakers, 3) words or phrases that respondents in Survey 1 had also underlined for strikingness, and 4) foreign vocabulary. I also mentioned that these categories were tentative and not exhaustive. Indeed, Venuti's descriptions of other successfully foreignising features may apply here. For instance, he discusses the combination of "current standard usage with colloquialisms and poetical archaisms" (Venuti [1995] 2008: 121). Where colloquialisms are concerned, this might explain why 'woke up in the dead of night' was underlined in Poem A, and why 'Drop dead, people!' was underlined four times in Poem D. The situation regarding archaisms is a bit more complicated: I included an anachronistic phrasing in Poem H ('The ugly few are we'), but no one underlined it. That said, 'make your lodging here' (Poem G) and 'take a room' (Poem C) both have a slightly old-fashioned ring, and respondents *did* underline those. Finally, three respondents underlined a string that could arguably be described as discontinuous at the level of syntax: this was 'help it had' in Poem A. One problem is: there are numerous overlaps with underlinings for strikingness in Survey 1. This suggests that an alternative approach may be useful in future to help disentangle (as much as possible) the phenomena of strikingness and foreignness. Furthermore, it is difficult to categorise several underlinings according to Venuti's criteria. **For this reason, these results lend only qualified support to his theory.**

The various ways of accounting for impressions of foreignness, as suggested by the four tentative categories above, restore an inner complexity to the phenomenology of 'the foreign'

366

that is obscured by binary models of translation. In such models, the ST author is either 'moved' toward the TT reader, or the TT reader is 'moved' toward the ST author. Yet even when these models celebrate the possibilities of cross-cultural exchange, there is often an alarming assumption of what we might call 'purity of origins'—that the domestic is always already *completely* domestic, and the foreign is always already *completely* foreign. Scholars are beginning to trouble these categories in consequential ways: taking a "polytemporal, polyspatial" (Ramazani 2020: 8) approach, Ramazani argues that today the "making of a poem, as of a pencil, amalgamates, reshapes, and compresses materials that span large swaths of the globe" (ibid.: 1). To put it another way: "If we were to submit almost any modern or contemporary poem to Ancestry.com, and if the units of analysis included allusions, techniques, etymologies, genres, forms, and rhetoric, the resulting pie chart, even if favoring one region in one era, would inevitably include others" (ibid.: 2). In short, as David Bellos (2020: n.p.) relates, "[t]here is no steel divider between what is domestic and what is not." However, what is now called for is not the abandonment of the concept of 'the foreign,' but a fuller investigation of the ways it participates in readers' engagements with translated texts, particularly in an era characterised by transnational flows of people and cultures. To that end, a methodology which gives more play to the triangulation of data (to try to determine *why* an item gives an impression of foreignness) will likely prove useful.

H_9 was the third and final hypothesis related to foreignisation. H_a predicted that a translation specifically undertaken to produce a foreignising effect (my translation of "Shunkan" [Poem H]) could score just as highly on the enjoyment scale as the existing translation of the same text. To determine this, I ran a Mann-Whitney U test to compare the relevant scores for Poems D and H. For this test, I combined enjoyment scores from both Survey 1 and 2. (Respondents in Survey 3 did not know the texts were translations.)

Results for H_9 showed:

- **whether respondents were analysed together as one big group, or as two separate groups (students and non-students), there was no significant difference between the enjoyment scores for Poems D and H** (according to a Mann-Whitney test)

  - *this is evidence in **favour** of the theory of foreignisation*

- **however, as we saw in H7, only one of the 'ostensibly foreignising' texts scored significantly higher than its counterpart for foreignisation score,** *and it was not Poem H***; this result is confirmed by a Mann-Whitney U test**

  - *this is evidence **against** the theory of foreignisation*

Results for this hypothesis are mixed: while my translation of "Shunkan" was just as well-received as the existing translation, there is still no evidence that it produced a stronger foreignising effect than Poem D. The results of the Mann-Whitney test for enjoyment scores follow. Since a significant result on this test would suggest that scores for one poem are significantly different than scores for the other, *the theory prefers the null hypothesis.* Thus, a *p*-value *above* 0.05 is evidence in favour of the theory (highlighted in green.)

Fig. 136.

Mann-Whitney U Test: Enjoyment Scores for Poems D and H (All Respondents)

	Group (Poem D or H)	N	Mean Rank	Sum of Ranks
Enjoyment Scores	PoemD	42	42.85	1799.50
	PoemH	39	39.01	1521.50
	Total	81		

Test Statistics[a]

	Enjoyment Scores
Mann-Whitney U	741.500
Wilcoxon W	1521.500
Z	-.737
Asymp. Sig. (2-tailed)	.461

a. Grouping Variable: Group (Poem D or H)

Fig. 137.

Mann-Whitney U Test: Enjoyment Scores for Poems D and H (Students)

	Group (Poem D or H)	N	Mean Rank	Sum of Ranks
Enjoyment Scores	PoemD	23	19.80	455.50
	PoemH	17	21.44	364.50
	Total	40		

Test Statistics[a]

	Enjoyment Scores
Mann-Whitney U	179.500
Wilcoxon W	455.500
Z	-.443
Asymp. Sig. (2-tailed)	.658
Exact Sig. [2*(1-tailed Sig.)]	.665[b]

a. Grouping Variable: Group (Poem D or H)

b. Not corrected for ties.

369

Fig. 138.

Mann-Whitney U Test: Enjoyment Scores for Poems D and H (Non-Students)

	Group (Poem D or H)	N	Mean Rank	Sum of Ranks
Enjoyment Scores	PoemD	19	22.21	422.00
	PoemH	22	19.95	439.00
	Total	41		

Test Statistics[a]

	Enjoyment Scores
Mann-Whitney U	186.000
Wilcoxon W	439.000
Z	-.605
Asymp. Sig. (2-tailed)	.545

a. Grouping Variable: Group (Poem D or H)

Meanwhile, the results of the Mann-Whitney test for foreignisation scores suggest there is no significant difference in foreignisation scores. The *p*-value there is *under* 0.05, and is thus highlighted in yellow. (Since Survey 1 respondents were not asked foreignisation questions, those surveys could not be included in this test.)

Fig. 139.

Mann-Whitney U Test: Foreignisation Scores (Survey 2--All Respondents)

	Group (Poem D or H)	N	Mean Rank	Sum of Ranks
Foreignisation Scores (Survey 2)	PoemD	12	9.79	117.50
	PoemH	9	12.61	113.50
	Total	21		

Test Statistics[a]

	Foreignisation Scores (Survey 2)
Mann-Whitney U	39.500
Wilcoxon W	117.500
Z	-1.049
Asymp. Sig. (2-tailed)	.294
Exact Sig. [2*(1-tailed Sig.)]	.310[b]

a. Grouping Variable: Group (Poem D or H)

b. Not corrected for ties.

So: my translation was enjoyed to about the same degree as the pre-existing translation, yet there is no good evidence that it produced a foreignising effect. What are the implications of this result? In Venuti's framework, the apparent lack of foreignising effect would presumably be explained in terms of an inappropriate response—perhaps a refusal on the part of readers to respond in terms of the translator's stated intentions (see e.g. Venuti 1998: 19). However, no stated intention accompanied my translation. (Indeed, it would not do to include material that might make it easier for respondents to guess the aims of my research.) Nevertheless, in the future that Venuti sees for literary translation, the onus would seem to fall on translators to *inform* readers how to correctly respond. As he writes, to "perceive foreignizing effects in a translation, a reader needs […] to be capable of articulating them as such" (Venuti 2008 [1995] 124). Who better to promote this new literacy than translators

themselves? As he writes more recently, "translators must be able to give a sophisticated account of their interpretive acts to the various readerships who rely on their work, both elite and popular, so as to support the development of a translation literacy" (Venuti 2019: 176). I cannot begrudge the desire for translators to be able to give 'sophisticated accounts' of their work. However, if this means insisting on a particular experience of the text as the most correct or appropriate one, this poses a distinct problem for the translator-who-also-researches-translation: one leaves oneself open to the criticism of interfering in the very responses one claims to be studying. Moreover, not every translator will want to be seen as (potentially) arbitrating or dictating responses.

One hypothesis (H_{10}) was tested in Survey 3. H_a predicted that respondents would evaluate the stimulus texts differently when the texts were presented simply as 'poems' as opposed to 'translations.' I investigated this with a Mann-Whitney U test on summed enjoyment scores. These were scores from two questions: 'Did you enjoy reading this poem?' and 'How enjoyable or aesthetically appealing did you find this poem?' I used these items because some Survey 3 respondents did not answer my question about 'literariness.' Scores from Survey 3, in which texts were presented simply as 'poems,' were compared with scores from Survey 2, in which the texts were presented as translations. For these tests, I did *not* combine scores from Surveys 1 and 2: given the small sample size for Survey 3, this would have resulted in comparing groups of highly unequal sizes, thereby reducing the power of the test to actually detect a difference. Moreover, in order to have enough text-as-poem responses, the criterion that respondents be native speakers of British English had to be dropped. (All respondents self-identified as native speakers of English, though some neglected to specify the variety of English.) Taking the exploratory nature of this survey into account, I set a slightly less restrictive significance threshold ($p = 0.10$).

Results for H_{10} show:

- **three poems (B, E and H), tended to have significantly higher enjoyment ratings when presented as translations from the Japanese as opposed to simply 'poems'** (according to a Mann-Whitney U test)

- **text condition was not associated with any significant changes in enjoyment scores for remaining 5 poems (A, C, D, F and G)** (according to a Mann-Whitney test)

In short, results for this hypothesis were rather mixed: we found *some* support for

the idea that foreknowledge of the translated condition of stimulus texts is associated with enjoyment scores. However, this was for only three texts, as the table below shows. In one sense, given the small sample size of Survey 3, to obtain even these significant results is noteworthy. At the same time, the question of the representativeness of the 'text-as-poem' respondents looms large. More research will have to be done in order to draw firmer conclusions.

Nevertheless, an interesting possibility suggests itself for B and E (both 'ostensibly domesticating' translations). If domestication produces the effect that readers are reading a non-translated text, then it might be the case that (under certain circumstances) knowledge of a text's foreign provenance may *distinguish* the literacy event in a way that enhances enjoyment. Perhaps for some readers this knowledge contributes to a sense of cosmopolitanism—that however similar the text may seem to a domestic, non-translated poem, one is broadening one's cultural horizons by reading a foreign text. Whether this cosmopolitanism is aware of the privilege that underwrites it is of course a crucial question (see e.g. Toth 2019). Nevertheless, foreknowledge of the translated-ness of the text may thus function as a kind of *pre-inscription of difference*, according to which readers make aesthetic allowances for the sameness of domesticating translation. By the same token, this pre-inscription of difference might also soften the blow of more deviant translation strategies: the other poem rated higher for enjoyment as a translation was H, an 'ostensibly foreignising' translation. It seems possible that the norm-breaking in H was so great that when readers knew the text was a translation, they could use this knowledge to explain away or make allowances for its deviations. With the text's eccentricities thus accounted for, readers may have felt less beholden to judge it according to domestic norms, hence the higher enjoyment scores for the text-as-translation condition.

Fig. 140. Mann Whitney U Test: Enjoyment Sums (H$_{10}$)

Poem	N$_1$ (Text-as-poem)	N$_2$ (Text-as-translation)	Mean Rank (Text-as-poem)	Mean Rank (Text-as-translation)	U value	p-value (exact sig. two-tailed)
A	4	12	5.50	9.50	12.00	0.170
B	4	12	3.63	10.13	4.50	0.013
C	4	12	7.88	8.71	21.50	0.770
D	4	12	7.38	8.88	19.50	0.599
E	3	9	3.17	7.61	3.50	0.064
F	3	9	5.67	6.78	11.00	0.727
G	3	9	5.17	6.94	9.50	0.482
H	3	9	2.83	7.72	2.50	0.036

5.3. SPEED READ SUMMARY (CONCLUSIONS AND EVIDENCE)

Fig. 141. Speed Read Summary

Hypothesis	Conclusions	Evidence	Support for Theory of FG?
H₁	Density of FG does not guarantee readers will perceive strikingness.	A binomial test revealed that only half of the texts have significantly more items underlined in FG than in BG lines.	No
	Without a weighting criterion, stylistic analyses based on the 'standard' model of FG *do not* accurately predict responses for strikingness.	A Spearman rank test revealed that predicted ranks of verse lines (based on amount of FG from stylistic analysis) do not agree with observed ranks (based on underlining activity of respondents) for most poems.	No
H₂	For most of the texts, literary expertise is not related to total amount of perceived FG.	A Mann-Whitney U test revealed no significant relationship between familiarity with theory and amount of FG identified per poem.	Yes (with caveat)
		A Kruskal-Wallis test revealed no significant correlation between literary training level and amount of FG identified per poem (for 7 out of 8 poems).	Yes (with caveat)
		Caveat: the number of items underlined per poem may have been artificially constrained by test instructions. That is, it may be that one subgroup *would* have underlined more items, were it not for the wording of the test instruction.	

	For some poems, expertise *is* implicated in amount of FG identified *per line.*	A Friedman test revealed qualitative differences in underlining activity *per line* between subgroups for some poems.	No
H₃	Poetry attitudes are not related to amount of FG identified per poem.	A Kruskal-Wallis test revealed no association between poetry attitudes and number of items underlined for strikingness.	Yes (with caveat)
		A Spearman rank test revealed no correlation between poetry attitudes and number of items underlined per poem.	Yes (with caveat)
		Caveat: the number of items underlined per poem may have been artificially constrained by test instructions. That is, it may be that one subgroup *would* have underlined more items, were it not for the wording of the test instruction.	
H₄	Age is not related to total amount of perceived FG per poem (when measured in number of items underlined for strikingness).	A Kruskal-Wallis test revealed no significant association between age group and number of words underlined for strikingness for 7 out of 8 poems.	Yes
	For some poems, older respondents were likely to indicate fewer instances of strikingness.	A Spearman rank test identified a moderate negative correlation between age and instances of strikingness for some poems.	No
	For some poems, older respondents were more likely	A Spearman rank test revealed a moderate	No

	to perceive longer units of text as striking in some poems.	positive correlation between age and length of underlining for some poems.	
	While older and younger respondents generally agreed on which *items* were striking, they disagreed on the relative strikingness of individual *verse lines*.	'Most Frequently Underlined Words' for both groups largely agree. However, there were sometimes large qualitative differences in trend lines for strikingness throughout texts.	Yes and No
H₅	Whether or not respondents were currently students did not seem to matter in terms of median number of words underlined per poem.	A Spearman rank test revealed a very strong positive correlation between the two groups in terms of median number of words underlined per poem. Caveat: the number of items underlined per poem may have been artificially constrained by test instructions. That is, it may be that one group *would* have underlined more items, were it not for the wording of the test instruction.	Yes (with caveat)
	Poem length also seems to predict how many items get underlined for strikingness.	A Spearman rank test indicated that poem length and amount of FG identified by stylistic analysis were highly correlated.	No
	Whether or not respondents were currently students *did* seem to matter in terms of which poems were found to be most literary.	A Spearman rank test found no correlation between the two groups' rankings of the texts for literariness.	No

		The amount of FG identified per text is not associated with literariness scores.	A Spearman rank test found no correlation between the amount of FG (as identified by stylistic analysis) and median literariness scores per poem.	No
		Number of striking items underlined and impression of literariness were less likely to be correlated for student respondents.	A Spearman rank test showed no significant correlations between these two variables for Poems A, B, C, E, G and H, for student respondents.	No
		Number of striking items underlined can be *negatively* correlated with impression of literariness.	A Spearman rank test found significant negative correlations between these two variables for Poems B, C, D, F and H (depending on the respondent group).	No
	H$_6$	Textual complexity (as measured by amount of FG) was not related to enjoyment.	A Spearman rank test identified no correlation between amount of FG (as identified by stylistic analysis) and enjoyment scores.	No
		Literariness and enjoyment go hand in hand for non-students, for these texts.	A Spearman rank test found strong positive correlations between scores for literariness and enjoyment for all texts for this group.	Yes
		Literariness and enjoyment do not necessarily go hand in hand for students, for these texts.	A Spearman rank test found positive correlations between scores for literariness and enjoyment for only half the texts for this group.	No
		Older respondents were less likely to enjoy texts that	A Mann-Whitney U test identified a significant	Unclear

	explicitly mentioned death or dying than younger respondents.	difference in enjoyment scores between student and non-student respondents for Poems A, D and H.	
H_7	The 'ostensibly foreignising' texts had higher median foreignisation scores than their competing translations.	This result is based on descriptive statistics. Caveat: only one 'ostensibly foreignising' text scored *statistically significantly* higher than its counterpart on median foreignisation score.	Yes (with caveat)
	Only one of the 'ostensibly foreignising' texts (Poem C) scored significantly higher than its counterpart (Poem G) for foreignisation score.	A rank biserial test revealed no significant correlation between text condition and median impression of foreignness scores for three out of four pairs.	No
	Within surveys, text condition was not systematically related to impression of foreignness.	A Friedman test revealed no systematic association between text condition and foreignisation scores within individual surveys.	No
	For every pair of competing translations but one (Poems A and E), the 'ostensibly foreignising' one has more variability in the foreignisation scores.	This is based on standard deviation of foreignisation scores.	No
H_8	Many of the items underlined by respondents can be categorised according to Venuti's descriptions of successfully foreignising features.	This is based on qualitative analysis of 'Most Frequently Underlined Words.' Caveat: it is difficult to categorise several underlined items according	Yes (with caveat)

		to Venuti's criteria	
H9	A translation specifically undertaken to be foreignising can be just as well-received (by this population of respondents).	A Mann-Whitney U test revealed no significant differences in enjoyment scores between Poems D and H.	Yes
	However, such a translation may not actually have a foreignising effect!	A Mann-Whitney U test revealed no significant differences in foreignisation scores between Poems D and H.	No
H10	Respondents tended to evaluate *some* texts more highly when presented as translations.	A Mann-Whitney U test revealed significant differences in enjoyment scores for Poems B, E and H across the two text conditions.	Unclear
Other	The 'aboutness' of the texts—as signalled in their titles—does not seem to diminish the strikingness of associated words in the text.	Words demoted from FG inventories due to their association with title words nevertheless appear in every list of 'Most Frequently Underlined Words,' apart from Sato's *Midaregami* no. 145.	N/A
Other	Strikingness may depend in part on intelligibility.	While some respondents indicated transliterated Japanese as striking, not a single respondent identified any item in Japanese script (kanji, hiragana) as striking, even when they had glosses for pronunciation or meaning.	N/A
Other	Enjoyment is not necessarily a component of 'literariness' for some readers	Results from a principal components analysis (PCA) suggest that the appreciation scale developed by Dixon et al. (1993) is not necessarily unidirectional	N/A

6. GENERAL CONCLUSIONS

This thesis has addressed two major questions:

1. *If the experience of reading is deeply subjective, then how do we account for the apparent lack of diversity of response reported in some empirical studies?*

2. *How does the foreknowledge of the translated-ness of texts participate in the affective and cognitive experiences of TL readers?*

Where the first question is concerned, I provided two answers. First: this apparent lack of diversity is sometimes due to researchers' own interpretations of their results. Second: overreliance on undergraduate respondents has not given diversity of response a fair chance to emerge. In expanding my pool of respondents to include people not normally included in this kind of research, my study has produced some good evidence in favour of the idea that patterns of response vary as a function of both textual and reader variables. Indeed, results for several hypotheses derived from the theory of FG suggest that the case for the text-directedness of poetry reading is not as strong as previously thought. Many of these results were found through comparing student and non-student responses. This suggests that in order to more fully understand the act of poetry reading, it is imperative that researchers in ESL begin to study *other types of readers* than the customary sample of undergraduates, however stratified by experience level they may seem. Moreover, as I suggested in Chapter 1, we should be careful when it comes to generalising beyond the text type investigated here: we lack the data to claim that what holds for poetry holds for other text types more generally. To that end, future ESL research might do well to divest from the poeticity-as-literariness model to gather data on other text types. According to the RSL survey (RSL 2017a: 12), poetry may

be the second most likely text type to be considered literature, but it is also one of the "least common types of reading matter" of British adults (ibid.: 10), alongside self-help books and comics or graphic novels.

While the answer to the second major question is far from settled, analysis of survey responses revealed a kind of divided subjectivity on the part of some respondents. Comments left by these readers suggested that, in the absence of an authority to (pre)validate their responses, they did not trust their personal interactions with the text to provide a solid basis for aesthetically evaluating poetry originating in a different cultural-linguistic milieu. I elaborated these findings in terms of Toth's (2019) notion of "the reading self" and the "self-in-the-world" to argue that such responses may indicate an 'ethics of alterity' at work: these readers' hesitations and concerns suggested a kind of demurral premised on the possibility of a more *culturally informed* response—to be instantiated at some future point, attendant upon further knowledge not just of the TTs but of the literary and cultural context of the STs. In such a phenomenology, no reading (or evaluation) is ever final, merely one in a series of possible readings—each reading presenting the subject with an opportunity to (better) recognise its own cultural-historical situatedness. Here, too, the question of text-type arises: if it is not uncommon, after Arnold's "The Study of Poetry" (1888), to encounter in Anglophone contexts the idea of poetry as the highest form of literature, then is it not possible that the weightiness of the genre may provoke a special unease when it comes to the activity of evaluation?

While Venuti's theory of foreignisation found only limited support in the statistical results, the experience of putting his theory to work—crossing the theory/practice divide—was (as documented in Chapter 3) both challenging and stimulating, exposing contradictions in the theory but disclosing new aesthetic possibilities at the same time. I found I could not make the crossing without identifying with the poet in some sense (cf. Venuti's [2008: 237-

239] disapproval of "simpatico"). Moreover, while I accept (and embrace) the likelihood that my TT would have "different meanings" for "different groups" (Venuti 1998: 12), I could not quite get myself to believe that my translation practice could ever "invent a minor language that cuts across cultural divisions and hierarchies" (ibid.: 13), addressing and serving multiple constituencies such that they all come to a newfound "recognition of translation" (ibid.: 13). In other words, I could not cross the theory/practice divide without ceding a certain illusion of control over response. (If Venuti invests *intentio translatoris* with the authority to arbitrate TT readers' responses, this seems closely tied to an uneasiness over the ability of translation features *themselves* to produce the hoped-for effects.) Nevertheless, attempting a foreignising translation presented me with the opportunity to stretch as a writer: the task seemed to require me to discover a new way of putting a text together, to integrate (SL) linguistic material I would not normally use into an aesthetically satisfying whole (even if it might only satisfy me in the end). This prompted a closer attentiveness to the feelings and sensations I rely upon to 'tell' me when a translation is finished, as well as the recognition that what I enjoy about translating poetry is partly the feeling of "being an embodied subject who is fully present to the moment" (Folkart 2007: 80). However, it was in partly *absenting* myself, in *not* allowing myself to translate as I usually do, that the resulting TT ultimately satisfied me. This recalls Yoshihara's writing method, in which "someone who is not Yoshihara Sachiko" sometimes comes along to contribute to the writing process (Yoshihara [1975] 1983: 195). Perhaps a similar sense of alterity—that it was not fully *me* in control—accounts for the strange way the TT affects me now. *It seems darker, denser than I remember.*

In the process of analysing survey results, this thesis furthermore suggested several methodological refinements, produced evidence that an important appreciation scale is not necessarily unidirectional, and generated further research questions in both theoretically and socially consequential research areas (e.g. this question from H$_6$: *Is emotional valence a*

better predictor of enjoyment for older readers than younger ones?).

Results suggest three main recommendations for future survey design:

1. Since we can no longer assume an additive model of literariness (i.e. literariness as a kind of 'sum' of FG features), it may be helpful to have respondents indicate the *degree to which* each underlined item contributes to their impression of literariness, and *why*.

2. To account for the 'performative' aspect of Likert scales (see H$_5$ results), it is important to allow readers to be 'divided subjects' capable of recognising where institutional values may conflict with their own; thus, it may help to include additional questions about how readers interpret contested terminology.

3. To avoid the charge that it is studies of literariness, in part, that are *keeping the idea of literariness in the world*, future methodologies might refrain from explicitly invoking the distinction between literary and non-literary texts, so as not to predispose respondents toward adopting the categories of the researcher.

While my surveys solicited comments on the final page, I am fortunate that some respondents also saw fit to provide additional information on other pages of their surveys, or in person, as this provided an opportunity to triangulate data (albeit in a non-systematic way). For instance, one respondent mentioned that they were unsure whether they would reply the same way on another day of the week, which highlighted for me the potential impact of respondents' initial affective states. If we understand triangulation to mean "the practice of using multiple sources of data or multiple approaches to analyzing data to enhance the credibility of a research study" (Hastings 2010: 1537), then we may point to the fact that

multiple statistical tests were run on the same data for many of the tested hypotheses. However, it is clear that future studies could incorporate more qualitative data sources like interviews or writing tasks to add nuance to the interpretation of quantitative data.

As with any research methodology, certain limitations must be addressed. There were three clear limitations with a potential impact on my findings: 1) the use of a non-random sample (coupled with low response rates); 2) the fact that most student surveys (all but three) were completed in a classroom setting; 3) the lack of questions relating to cultural familiarity with Japan. The first limitation exists because it was not feasible to repeatedly contact members of my original random sample until the desired number of responses was reached. (I discussed the implications of this in section 3.3.) The second limitation exists because the initial survey instruction ("Please do this survey at a time and place where you usually read for pleasure") resulted in a very low response rate. The potential impact of this limitation is that the institutional setting of the activity may have primed students to respond in ways less reflective of their own individual reading practices and more reflective of practices associated with the particular classroom/course. The third limitation exists because I was not aware of a pre-existing cultural familiarity scale that would account for all potentially relevant dimensions of familiarity, and I lacked the time and resources to independently develop and validate such a scale. One consequence of this limitation is that we lack the ability to account for the potential effects of such background knowledge on foreignisation scores. Overcoming the first limitation in future research will depend in part on the existence of (and researchers' access to) lists of target populations; it may also depend on implementing more effective means of reducing non-response (see e.g. Dillman et al. 2014 for more on 'mixed mode' surveys employing social exchange theory). Overcoming the second limitation may also require a more sophisticated data collection model to better motivate students to complete the surveys according to the original instruction. Where the third limitation is concerned, it may

serve to ask whether respondents have ever studied Japanese language or culture, and to gauge the frequency with which they watch, read or listen to media produced in Japan.[59]

It is hoped that implementing these suggestions will improve the quality of research in this fascinating area, and that ESL researchers might soon begin in earnest to connect their findings about reader behaviour to ongoing research in socially consequential areas like literacy, creative arts therapy, gerontology, library science and the medical humanities. One potential future line of inquiry that springs out of this study concerns 'depth of processing'— this refers to "the degree to which a stimulus is processed at different levels of mind which affects the likelihood of its being remembered" (Matsumoto 2009: 156). While the originators of the theory were primarily concerned with readers' memory of individual lexical items (see Craik and Lockhart 1972), other researchers have extended the idea toward text comprehension (see e.g. Sanford and Sturt 2002). If for some texts, older readers tend to find longer units of text striking, it seems intuitively correct to hypothesise that these readers are paying attention to levels of meaning higher than lexis. One task for the future is to determine why: are younger readers adopting a faster, shallower reading style due to claims on their attention? Does their screen usage have a bleed-over effect on the reading styles they bring to paper-based environments (see Delgado et al. 2018)?

Finally, it is hoped that the publication of more research in ESL will contribute to an atmosphere in which humanistic and empirical study can enrich one another. Here I side with Van Peer, Hakemulder and Zyngier (2012: 29) when they write: "If there is one thing that the

---

[59] Related to this is the Japanese reputation of the source authors. It would be no exaggeration to say that Tanikawa, Ishigaki and Akiko have enjoyed a critical and popular reception in Japan that distinguishes them from their lesser-known contemporaries. Given their distinctive styles and concerns, it seems sensible to wonder what impact this particular selection of poets might have had on results. This is difficult to say; in a sense, it could only be answered by presenting the same respondents with another set of texts by a different set of authors, such that the 'new' authors differed from the 'old' in as few respects as possible save one: their styles and concerns would have to be less distinctive. But this opens a new can of methodological worms: on what basis would we determine this *anti*-distinctiveness? Moreover, since respondents would be evaluating TTs, the resulting differences in response might be explained away by differences in translators or translation strategies applied to the 'new' group of authors.

scientific study of culture reveals again and again, it is the richness and variety of the world, not only the variation between individuals and groups, but also—often—the patterns and regularities that are there under a surface of variation."

7. BIBLIOGRAPHY

6 [sic], Perri and Christine Bellamy. *Principles of Methodology: Research Design in Social Science*. Los Angeles: SAGE: 2012.

Adler, Emily Stier and Roger Clark. *An Invitation to Social Research*. 5th ed. Stamford, CT: Cengage Learning, 2015.

Agirrezabal, Manex, Aitzol Astigarraga, Bertol Arrieta and Mans Hulden. "ZeuScansion: A Tool for Scansion of English Poetry." *Journal of Language Modelling* 4, no. 1 (2016): 3–28.

Alters, Sandra. *Biology: Understanding Life*. 3rd ed. Boston, MA: Jones and Bartlett, 2000.

Anthony, Laurence. *AntConc* (Version 3.5.7). Tokyo: Waseda University, 2018.

Anttila, Arto. "Stress, Phrasing, and Auxiliary Contraction in English." In *The Morphosyntax-Phonology Connection: Locality and Directionality at the Interface*, edited by Vera Gribanova and Stephanie S. Shih, 143-170. Oxford: Oxford University Press, 2017.

Arnold, Matthew. "The Study of Poetry." In *Essays in Criticism: Second Series*, 1-55. London: Macmillan, 1888.

Arntzen, Sonja. Review of *Japanese Women Poets: An Anthology*, by Hiroaki Sato. *Japan Forum* 21, no. 1: 155-157.

Aron, Arthur, Helen E. Fisher and Greg Strong. "Romantic Love." In *The Cambridge Handbook of Personal Relationships*, edited by Anita L. Vangelisti and Daniel Perlman, 595–614. Cambridge: Cambridge University Press, 2006.

Arrojo, Rosemary. "The 'Death' of the Author and the Limits of the Translator's Visibility." In *Translation as Intercultural Communication*, edited by Mary Snell-Hornby, Zuzana Jettmarová and Klaus Kaindl, 21-32. Amsterdam: John Benjamins, 1997.

Askin, Ridvan. "Objects." In *The Cambridge Companion to Literature and the Posthuman*, edited by Bruce Clarke and Manuela Rossini, 170-181. Cambridge: Cambridge University Press, 2017.

Atwell, John E. *Ends and Principles In Kant's Moral Thought*. Dordrecht: Martinus Nijhoff Publishers, 1986.

Austin, J.L. *How to Do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press, 1975.

Aviram, Amittai F. *Telling Rhythm: Body and Meaning in Poetry*. Ann Arbor: University of Michigan Press, 1994.

Babbie, Earl. *Survey Research Methods*. 2nd ed. Belmont, CA: Wadsworth, 1990.

Backus, Megan, trans. *Kitchen*. By Banana Yoshimoto. New York: Grove Press, 1993.

Baichiman, Janīn [Janine Beichman]. "*Oriori-no uta*-no eiyaku" [the English translation of (Ōoka Makoto's) *Oriori-no uta*]. *Ōoka Makoto kenkyū* 3 (2017): 72-95.

Baker, Mona. "Reframing Conflict in Translation." *Social Semiotics* 17, no. 2 (2007): 151-169.

Baldick, Chris. *Criticism and Literary Theory 1890 to the Present*. London: Routledge, 1996.

Baron, Naomi S. "Commas and Canaries: the Role of Punctuation in Speech and Writing." *Language Sciences* 23, no. 1 (2001): 15-67.

Barthes, Roland. *The Preparation of the Novel*. 1978-1980. Translated by Kate Briggs. New York: Columbia University Press, 2011.

Barthes, Roland. *Writing Degree Zero*. 1953. Translated by Annette Lavers and Colin Smith. London: Jonathan Cape, 1967.

Baynham, Mike and Tong King Lee. *Translation and Translanguaging*. Oxon: Routledge, 2019.

BBC Front Row. Interview with Junko Takekawa and Nicole Coolidge Rousmaniere.

London: BBC Radio 4, 17 September 2019. Available online at:

https://www.bbc.co.uk/sounds/play/m0008hxz.

Beichman, Janine. 2002a. *Embracing the Firebird: Yosano Akiko and the Birth of the Female Voice in Modern Japanese Poetry*. Honolulu: University of Hawai'i Press, 2002.

Beichman, Janine, trans. 2002b. *Midaregami* no. 145, by Yosano Akiko. In *Embracing the Firebird: Yosano Akiko and the Birth of the Female Voice in Modern Japanese Poetry*, 214. Honolulu: University of Hawai'i Press, 2002.

Beichman, Janine, trans. "Shijimi Clams," by Rin Ishigaki. In *The Columbia Anthology of Modern Japanese Literature*, vol. 2, edited by J. Thomas Rimer and Van C. Gessel, 415. New York: Columbia University Press, 2007.

Belfi, Amy M., Edward A. Vessel and G. Gabrielle Starr. "Individual Ratings of Vividness Predict Aesthetic Appeal in Poetry." *Psychology of Aesthetics, Creativity, and the Arts* 12, no. 3 (2018): 341-350.

Bellos, David. "The Myths and Mysteries of Literary Translation." The W.G. Sebald Lecture 2020. Delivered online, 29 June 2020. Available online at: https://www.youtube.com/watch?v=GnvC8ufhnt8.

Bennett, Karen. 2019a. "TEALS: Translationality." Accessed 13 October 2020. https://www.cetaps.com/teals-translationality/.

Bennett, Karen. 2019b. Review of *Translationality: Essays in the Translational-Medical Humanities*, by Douglas Robinson, *The Translator* 25, no. 2 (2019): 174-176.

Benthien, Claudia, Jordis Lau and Maraike M. Marxsen. *The Literariness of Media Art*. New York: Routledge, 2019.

Berg, Mari-Ann. Review of *Literature*, by Peter Widdowson, *Language and Literature* 9, no. 3 (2000): 281-283.

Berlina, Alexandra, trans. *Viktor Shklovsky: A Reader*. New York: Bloomsbury, 2017.

Bernstein, Charles. *Content's Dream: Essays 1975–1984*. Los Angeles, CA: Sun & Moon, 1986.

Bertens, Hans. *Literary Theory: The Basics*. London: Routledge, 2001.

Bevitori, Cinzia. "Discursive Constructions of the Environment in American Presidential Speeches 1960–2013: A Diachronic Corpus-Assisted Study." In *Corpora and Discourse Studies*, edited by Paul Baker and Tony McEnery, 110-133. New York: Palgrave, 2015.

Birnbaum, Alfred, trans. *All She Was Worth*, by Miyuki Miyabe. Tokyo: Kodansha, 1996.

Blair, Johnny, Ronald F. Czaja and Edward A. Blair. *Designing Surveys: A Guide to Decisions and Procedures*. 3rd ed. Los Angeles: SAGE, 2014.

Bliss, Frank W. and Earl R. MacCormac. "Deep Structures in Poetry." *Poetics and Theory of Literature* 2, no. 2. (1977): 227-240.

Blyth, R. H. *Haiku*, vol. 1. 1949. Tokyo: The Hokuseido Press, 1981.

Boase-Beier, Jean. *A Critical Introduction to Translation Studies*. London: Continuum, 2011.

Boase-Beier, Jean. *Stylistic Approaches to Translation.* Manchester: St. Jerome, 2006.

Boase-Beier, Jean. *Translation and Style*. 2nd ed. London: Routledge, 2020.

Boase-Beier, Jean. "Who Needs Theory?" In *Translation: Theory and Practice in Dialogue*, edited by Antoinette Fawcett, Karla L. Guadarrama García, and Rebecca Hyde Parker, 25-38. London: Continuum, 2010.

Bornstein, Diane D. *An Introduction to Transformational Grammar*. Lanham: University Press of America, 1984.

Bortolussi, Marisa and Peter Dixon. "The Effects of Formal Training on Literary Reception." *Poetics* 23, no. 6 (1996): 471-487.

Bortolussi, Marisa and Peter Dixon. *Psychonarratology: Foundations for the Empirical Study of Literary Response.* Cambridge: Cambridge University Press, 2003.

Bourdieu, Pierre. *Distinction: A Social Critique of the Judgment of Taste*. Translated by

    Richard Nice. Cambridge, MA: Harvard University Press, 1984.

Bourdieu, Pierre. "Outline of a Sociological Theory of Art Perception." 1968. In *The Field of*

    Cultural Production: Essays on Art and Literature, edited by Randal Johnson, 215-

    237. New York: Columbia University Press, 1993. No translator noted.

Bownas, Geoffrey and Anthony Thwaite, trans. "The Isolation of Two Milliard Light Years,"

    by Tanikawa Shuntarō. In *The Penguin Book of Japanese Verse*, translated by

    Geoffrey Bownas and Anthony Thwaite. Middlesex: Penguin, 1964.

Brazell, Karen. Review of *From the Country of Eight Islands*, edited by Hiroaki Sato and

    Burton Watson, *The Journal of Asian Studies* 42, no. 2 (1983): 417-419.

Bright, William. "Poetic Structure in Oral Narrative." In *Spoken and Written Language:*

    Exploring Orality and Literary, edited by Deborah Tannen, 171-184. Norwood, NJ:

    Ablex, 1982.

Brogan, T.V.F. "Ictus." In *The Princeton Encyclopedia of Poetry and Poetics*, 4th ed.,

    edited by Roland Greene, Stephen Cushman, Clare Cavanagh, Jahan Ramazani and

    Paul Rouzer, 657-658. Princeton: Princeton University Press, 2012.

Brooks, Larry W., Donald F. Dansereau, Joni E. Spurlin and Charles D. Holley. "Effects of

    Headings on Text Processing." *Journal of Educational Psychology* 75, no. 2 (1983):

    292-302.

Brouillette, Sarah. *Postcolonial Writers in the Global Literary Marketplace*. London:

    Palgrave Macmillan, 2007.

Brower, Robert H. and Earl Miner, trans. *Fujiwara Teika's 'Superior Poems of Our Time.'*

    Stanford: Stanford University Press, 1967.

Bruhn, Jørgen and Anne Gjelsvik. "David Simon's Novel Cop Show." *New Review of Film*

    and Television Studies 11, no. 2 (2013): 133-153.

Burke, Michael and Kristy Evers. "Formalist Stylistics." In *The Routledge Handbook of Stylistics*, edited by Michael Burke, 31-44. London: Routledge, 2014.

Cantos Gómez, Pascual. *Statistical Methods in Language and Linguistic Research*. Sheffield: Equinox, 2013.

Carnie, Andrew. *Syntax: A Generative Introduction*. 3rd ed. Hoboken: Wiley-Blackwell, 2013.

Carter, Roland and Walter Nash. "Language and Literariness." *Prose Studies* 6, no. 2 (1983): 123-141.

Chomsky, Noam. 1964a. "Current Issues in Linguistic Theory." In *The Structure of Language: Readings in the Philosophy of Language*, edited by Jerry A. Fodor and Jerrold J. Katz, 50-118. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Chomsky, Noam. 1964b. "Degrees of Grammaticalness." In *The Structure of Language: Readings in the Philosophy of Language*, edited by Jerry A. Fodor and Jerrold J. Katz, 384-389. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Chomsky, Noam. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: The MIT Press, 1988.

Chomsky, Noam. *Rules and Representations*. New York: Columbia University Press, 1980.

Cirillo, Robert. "A Fresh Look at the Debate on Floating Quantifiers." *Language and Linguistics Compass* 6, no. 12 (2012): 796-815.

CLiPS (Computational Linguistics & Psycholinguistics Research Center). 2018. "Penn Treebank II Tag Set." https://www.clips.uantwerpen.be/pages/mbsp-tags. Accessed 27 May, 2019.

Cohen, Jacob. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

Colangelo, Stefano. "Haiku e altre piccole tempeste" [haiku and other little storms]. In

Culture allo specchio: Arte, letteratura, spettacolo e società tra il Giappone e l'Europa, edited by Wada Tadahiko and Stefano Colangelo, 116-129. Bologna: I Libri di Emil, 2012.

Cole, Stephen. "Why Sociology Doesn't Make Progress like the Natural Sciences." *Sociological Forum* 9, no. 2 (1994): 133-154.

Collins, Peter. "The Progressive Aspect in World Englishes: A Corpus-based Study." *Australian Journal of Linguistics* 28, no. 2 (2008): 225-249.

Constable, John. "A Pilot Study in the Epidemiology of Representations." *Human Nature* 8, no. 2 (1997): 171-203.

Cooke, Raymond. *Velimir Khlebnikov: A Critical Study*. Cambridge: Cambridge University Press, 1987.

Craik, Fergus I. M. and Robert S. Lockhart. "Levels of Processing: A Framework for Memory Research." *Journal of Verbal Learning and Verbal Behavior* 11, no. 6 (1972): 671-684.

Creaghead, Nancy A. and Kenneth G. Donnelly. "Comprehension of superordinate and subordinate information by good and poor readers." *Language, Speech, and Hearing Services in Schools* 13, no. 3 (1982): 177-186.

Creswell, John W. *Research Design*. 4th ed. Los Angeles: Sage, 2014.

Croft, William and D. Alan Cruse. *Cognitive Linguistics*. Cambridge: Cambridge University Press, 2004.

Cronk, Nicholas. "Aristotle, Horace, and Longinus: the Conception of Reader Response." In *The Cambridge History of Literary Criticism*, vol. III, edited by Glyn P. Norton, 199-204.

Crystal, David. *The Cambridge Encyclopedia of the English Language.* 2nd ed. Cambridge: Cambridge University Press, 2003.

Crystal, David. *A Dictionary of Linguistics and Phonetics*. 6th ed. Oxford: Blackwell, 2008.

Culler, Jonathan. *Literary Theory: A Very Short Introduction*. 1st ed. Oxford: Oxford University Press, 1997.

Culler, Jonathan. *Literary Theory: A Very Short Introduction*. 2nd ed. Oxford: Oxford University Press, 2011.

Culler, Jonathan. *Structuralist Poetics: Structuralism, Linguistics and the Study of Literature*. London: Routledge, 1975.

Culpeper, Jonathan. "Keyness: Words, Parts-of-Speech and Semantic Categories in the Character-Talk of Shakespeare's *Romeo and Juliet." International Journal of Corpus Linguistics* 14, no. 1 (2009): 29-59.

Cutler, Anne. "Lexical Stress in English Pronunciation." In *The Handbook of English Pronunciation*, edited by Marnie Reed and John Levis, 106-124. Hoboken: Wiley, 2015.

Czarniawska, Barbara. *Narrating the Organization: Dramas of Institutional Identity*. Chicago: University of Chicago Press, 1997.

D'Alessandro, Roberta and Marc van Oostendorp. "When Imperfections Are Perfect: Prosody, Phi-features and Deixis in Central and Southern Italian Vocatives." In *Romance Languages and Linguistic Theory 10: Selected Papers from 'Going Romance' 28, Lisbon*, edited by Ernestina Carrilho, Alexandra Fiéis, Maria Lobo and Sandra Pereira, 61-82. Amsterdam: John Benjamins, 2016.

Davies, Mark. *BYU-BNC*. The British National Corpus. 2004. Available online at: https://corpus.byu.edu/bnc/.

De Beaugrande, Robert and Wolfgang U. Dressler. *Introduction to Text Linguistics*, translated by Robert de Beaugrande. London: Longman, 1981.

De Certeau, Michel. "Reading as Poaching." In *The Practice of Everyday Life*, translated by

Steven F. Rendall, 165-176. Berkeley: University of California Press, 1984.

Delgado, Pablo, Cristina Vargas, Rakefet Ackerman, and Ladislao Salmerón. "Don't throw Away Your Printed Books: A Meta-Analysis on the Effects of Reading Media on Reading Comprehension." *Educational Research Review* 25 (2018): 23-38.

Díaz-Diocaretz, Myriam. *Translating Poetic Discourse: Questions on Feminist Strategies in Adrienne Rich*. Amsterdam: John Benjamins, 1985.

Dillman, Don A., Jolene D. Smyth and Leah Melani Christian. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. 4th ed. Hoboken: Wiley, 2014.

Dixon, Peter, Marisa Bortolussi, Leslie C. Twilley, and Alice Leung. "Literary Processing and Interpretation: Towards Empirical Foundations." *Poetics* 22, no. 1 (1993): 5-33.

Dunning, Ted. "Accurate Methods for the Statistics of Surprise and Coincidence." *Computational Linguistics* 19, no. 1 (1993): 61-74.

Easthope, Antony. *Literary into Cultural Studies*. London: Routledge, 1991.

Easthope, Antony. *Poetry as Discourse*. London: Methuen, 1983.

Eco, Umberto. "Interpretation and History." In *Interpretation and Overinterpretation*, edited by Stefan Collini, 23-44. Cambridge: Cambridge University Press, 1992.

Eco, Umberto. *The Limits of Interpretation*. Bloomington: Indiana University Press, 1990.

Eco, Umberto. "Towards a Semiotic Inquiry into the Television Message." 1965. Translated by Paola Splendore. In *Television: Critical Concepts in Media and Cultural Studies*, vol. 2, edited by Toby Miller, 3-19. London: Routledge, 2003.

Eco, Umberto. *Trattato di Semiotica Generale*. Milan: Bompiani, 1975.

Eichenbaum, Boris. "The Theory of the 'Formal Method.'" 1926. In *Russian Formalist Criticism: Four Essays*, translated and with an introduction by Lee T. Lemon and Marion J. Reis, 99-139. Lincoln: University of Nebraska Press, 1965.

Elliott, William I. and Kazuo Kawamura, trans. "Two Billion Light-Years of Solitude." In *Shuntarō Tanikawa: Selected Poems*, translated by William I. Elliott and Kazuo Kawamura, 17-18. London, Carcanet: 1998.

Elsness, Johan. *The Perfect and the Preterite in Contemporary and Earlier English*. Berlin: Mouton de Gruyter, 1997.

Emmott, Catherine, Anthony J. Sanford, and Lorna I. Morrow. "Capturing the Attention of Readers? Stylistic and Psychological Perspectives on the Use and Effect of Text Fragmentation in Narratives." *Journal of Literary Semantics* 35, no. 1 (2006): 1-30.

Erlich, Victor. *Russian Formalism: History – Doctrine*. 1st ed. The Hauge: Mouton, 1965.

Erlich, Victor. *Russian Formalism: History – Doctrine*. 4th ed. The Hauge: Mouton, 1980.

Firth, J. R. "Modes of Meaning." In *Papers in Linguistics 1934–1951*, edited by J. R. Firth 190–215. London: OUP.

Firth, J. R. "A Synopsis of Linguistic Theory 1930-55." In *Selected Papers of J. R. Firth 1952–1957*, edited by F. R. Palmer. 168–205. London: Longmans, 1957.

Fish, Stanley. 1976. "Interpreting the *Variorum*." In *Is There a Text in This Class? The Authority of Interpretive Communities*, 147-173. Cambridge, MA: Harvard University Press, 1980.

Fish, Stanley. 1980a. "Introduction." In *Is There a Text in This Class? The Authority of Interpretive Communities*, 1-17. Cambridge, MA: Harvard University Press, 1980.

Fish, Stanley. 1980b. *Is There a Text in This Class? The Authority of Interpretive Communities*. Cambridge, MA: Harvard University Press, 1980.

Fish, Stanley. "Literature in the Reader: Affective Stylistics." 1970. In *Is There a Text in This Class? The Authority of Interpretive Communities*, 21-67. Cambridge, MA: Harvard University Press, 1980.

Fish, Stanley. "What Is Stylistics and Why Are They Saying Such Terrible Things About It?"
1973. In *Is There a Text in This Class? The Authority of Interpretive Communities*,
68-96. Cambridge, MA: Harvard University Press, 1980.

Flores Jurado, Julieta. "Pearls, Diamonds and Coins: Fashioning the Beloved's Body in Carol
Ann Duffy's Poetry." *Lectora* 26 (2020): 33-48.

Folkart, Barbara. *Second Finding: A Poetics of Translation.* Ottawa: University of Ottawa
Press, 2007.

Fónagy, I. and Magdics K. "Speed of Utterance in Phrases of Different Lengths." *Language
and Speech* 3 (1960): 179-192.

Fowler, Roger. *Linguistic Criticism.* 2nd ed. Oxford: Oxford University Press: 1996.

Fowler, Roger. 'Literature.' In *Encyclopedia of Literature and Criticism*, edited by Martin
Coyle, Peter Garside, Malcolm Kelsall and John Peck, 3-26. London and New York:
Routledge, 1990.

Fowler, Roger. "Polyphony in *Hard Times*." In *Language, Discourse, and Literature*, edited
by Ronald Carter and Paul Simpson, 77–93. London: Allen and Unwin, 1989.

Frölich, Grit. "Culture and Invention—Umberto Eco and the Aesthetics of Luigi Pareyson."
In *Umberto Eco in His Own Words*, edited by Torkild Thellefsen and Bent Sørensen,
34-40. Boston: De Gruyter, 2017.

Fry, D. B. "The Frequency of Occurrence of Speech Sounds in Southern English." *Archives
Néerlandaises de Phonétique Expérimentale* 20 (1947): 103-106.

Fry, Paul H. *Theory of Literature*. New Haven: Yale University Press, 2012.

Gandón Chapela, Evelyn. "A Corpus-Based Analysis of Invisible Elements in English: A
Pilot-Study on Verb-Governed Ellipsis." In *Reverberations of Silence: Silenced Texts,
Sub-Texts and Authors in Literature, Language and Translation*, edited by Márta
Pellérdi and Gabriella Reuss, 281-308. Newcastle upon Tyne: Cambridge Scholars

Publishing, 2013.

Gao, Xin. "Nonparametric Statistics." In *Encyclopedia of Research Design*, edited by Neil J.
Salkind, 915-920. Los Angeles: SAGE, 2010.

Gardner, Dee and Mark Davies. "Pointing Out Frequent Phrasal Verbs: A Corpus-Based
Analysis." *TESOL Quarterly* 41, no. 2 (2007): 339-359.

Garson, Justin. "Emergence." In *The Philosophy of Science: An Encyclopedia*, vol. 1, edited
by Sahotra Sarkar and Jessica Pfeifer, 230-235. New York: Routledge, 2006.

Garvin, Paul L., trans. *A Prague School Reader on Esthetics, Literary Structure, and Style*.
Washington, D.C.: Georgetown University Press, 1964.

Garza, James. "Dominant Forms and Marginal Translations: Re-reading the Emergence of
the Novel in Nineteenth-Century Japan." *Comparative Critical Studies* 17, no. 3
(2020): 413-432.

Ghasemi, Asghar and Saleh Zahediasl "Normality Tests for Statistical Analysis: A Guide for
Non-Statisticians." *International Journal of Endocrinology and Metabolism* 10, no. 2
(2012): 486–489.

Gillen, Julia and Roger Hancock. "Eating." In *International Perspectives on Early Childhood
Research*, edited by Julia Gillen and Catherine Ann Cameron, 100-113. New York:
Palgrave Macmillan, 2010.

Gimson, A. C. *An Introduction to the Pronunciation of English*. London: Arnold, 1962.

Goldston, M. Jenice and Laura Downey. *Your Science Classroom: Becoming an Elementary /
Middle School Science Teacher*. Los Angeles: Sage, 2013.

Goodwyn, Andrew. "The Status of Literature: English Teaching and The Condition of
Literature Teaching In Schools." *English in Education* 46, no. 3 (2012): 212-227.

Goodwyn, Andrew and Kate Findlay. "I. A. Richards and the Paradox of Practical Criticism:
A Response to David W. West." *Changing English* 10, no. 1 (2003): 101-106.

Gordon, Peter C., Matthew W. Lowder and Renske S. Hoedemaker. "Reading in Normally

    Aging Adults." In *Cognition, Language and Aging*, edited by Heather Harris Wright,

    165–191. Amsterdam: John Benjamins.

Graham, Gordon. *Theories of Ethics: An Introduction to Moral Philosophy with a Selection

    of Classic Readings*. New York: Routledge, 2011.

Gravetter, Frederick J. and Larry B. Wallnau. *Essentials of Statistics for the Behavioral

    Sciences*. 4th ed. Pacific Grove, CA: Wadsworth, 2002.

Grice, H. P. "Logic and Conversation." In *Syntax and Semantics 3: Speech Acts*, edited by P.

    Cole and J. L. Morgan, 41-58. New York: Academic Press, 1975.

Grimshaw, Gina M., Ana Adelstein, M. Philip Brydenand G. E. MacKinnon. "First-Language

    Acquisition in Adolescence: Evidence for a Critical Period for Verbal Language

    Development." *Brain and Language*, 63, no. 2 (1998): 237-255.

Grotjahn, R. and G. Altmann. "Modelling the Distribution of Word Length: Some

    Methodological Problems." In *Contributions to Quantitative Linguistics*, edited by

    Reinhard Köhler and Burghard B. Rieger, 141-154. Dordrecht: Springer, 1993.

Gutierrez, Kris D. "Unpackaging Academic Discourse." *Discourse Processes* 19, no. 1

    (1995): 21-37.

Hakemulder, Jèmeljan F. "Foregrounding and Its Effect on Readers' Perception." *Discourse

    Processes* 38, no. 2 (2004): 193-218.

Hall, John. "Cross-Sectional Survey Design." In *Encyclopedia of Survey Research Methods*,

    edited by Paul J. Lavrakas. Thousand Oaks: SAGE Publications, 2008.

Hall, Stuart. "Encoding/Decoding." 1980. In *Culture, Media, Language*, edited by Stuart

    Hall, Dorothy Hobson, Andrew Lowe and Paul Willis, 107-116. London: Taylor &

    Francis, 2005.

Halliday, M. A. K. "Linguistic Function and Literary Style: An Inquiry into the Language of

William Golding's *The Inheritors*." In *Literary Style: A Symposium*, edited by Seymour Chatman, 330-368. London: Oxford University Press, 1971.

Hanauer, David. "Integration of Phonetic and Graphic Features in Poetic Text Categorization Judgements." *Poetics* 23, no. 5 (1996): 363-380.

Hanauer, David. "Reading Poetry: An Empirical Investigation of Formalist, Stylistic, and Conventionalist Claims." *Poetics Today* 19, no. 4 (1998): 565-580.

Hanssen, Beatrice. "'Dichtermut' and 'Blödigkeit': Two Poems by Friedrich Hölderlin, Interpreted by Walter Benjamin." In *Walter Benjamin and Romanticism*, edited by Beatrice Hanssen and Andrew Benjamin, 139-162. London: Continuum, 2003.

Harries, Phillip. Review of *From the Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, *The Journal of the Association of Teachers of Japanese* 17, no. 1 (1982): 98-101.

Harvey, Rhonda L. and Donnie Dutton. "Reading Interests of Older Adults." *Educational Gerontology* 4, no. 3 (1979): 209-214.

Hastings, Sarah L. "Triangulation." In *Encyclopedia of Research Design,* vol. III, edited by Neil J. Salkind, 1537-1540. Thousand Oaks: SAGE Publications, 2010.

Havránek, Bohuslav. "The Functional Differentiation of the Standard Language." 1932. In A *Prague School Reader on Esthetics, Literary Structure, and Style*, selected and translated by Paul L. Garvin, 3-16. Washington, D.C.: Georgetown University Press, 1964.

Heath, Shirley Brice. 1982a. "Protean Shapes in Literacy Events: Ever-Shifting Oral and Literate Traditions." In *Spoken and Written Language: Exploring Orality and Literary*, edited by Deborah Tannen, 91-117. Norwood, NJ: Ablex, 1982.

Heath, Shirley Brice. 1982b. "What No Bedtime Story Means: Narrative Skills at Home and School." *Language in Society* 11, no. 1. (1982): 49-76.

Heilbrun, Carolyn G. *Writing a Woman's Life*. New York: Ballantine Books, 1989.

Hirshfield, Jane with Aratani Mariko. *The Ink Dark Moon: Love Poems by Ono no Komachi and Izumi Shikibu*. New York: Charles Scribners' Sons, 1986.

Hoffstaedter, Petra. "Poetic Text Processing and Its Empirical Investigation." *Poetics* 16, no. 1 (1987): 75-91.

Hogan, P. C. "Literature, God, & the Unbearable Solitude of Consciousness." *Journal of Consciousness Studies* 11, nos. 5/6 (2004): 116-142.

Huddleston, Rodney and Geoffrey K. Pullum, eds. *The Cambridge Grammar of the English Language.* Cambridge: Cambridge University Press, 2002.

IBM Corp. IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp, 2017.

Ijiri, Kayoko. "Kokusai haiku-to kigo: Aruzenchin haiku-wo megutte" [International haiku and season words: on Argentine haiku]. *Kyōto Sangyō Daigaku Ronshū Jinbun Kagaku Keiretsu* 45 (2012): 315-331.

Isakson, Carol S. and Jan H. Spyridakis. "The Influence of Semantics and Syntax on What Readers Remember." *Technical Communication* 50, no. 4 (2003): 538-553.

Ishigaki, Rin. "Nenpu, sankō bunken" [chronology, bibliography]. In *Ishigaki Rin shishū* [selected poems of Ishigaki Rin], 247-254. Tokyo: Haruki Bunko, 1998.

Ishigaki, Rin. "Shijimi" [clam]. 1968. In *Ishigaki Rin shishū* [selected poems of Ishigaki Rin], 67. Tokyo: Haruki Bunko, 1998.

Itō, Hiromi. "Rō-no-yō-na kubi-kara kyūkana-no ha-gyō" ['h' sounds in historical kana from a wax-like neck]. 1981. Folding insert included in *Yoshihara Sachiko zenshi III* [collected poems of Yoshihara Sachiko volume 3], 3-6. Tokyo: Shichōsha, 2012.

Jacob, Pierre. "Intentionality." In *Stanford Encyclopedia of Philosophy*. Article published Aug. 7, 2003. Available online at: https://plato.stanford.edu/entries/intentionality/.

Jakobson, Roman. "Closing Statement: Linguistics and Poetics." In *Style in Language*, edited

by Thomas A. Seobok, 350-377. Cambridge, MA: The M.I.T. Press, 1960.

Jakobson, Roman. 1935. "Marginal Notes on the Prose of the Poet Pasternak." In *Language in Literature*, edited by Krystyna Pomorska and Stephen Rudy, 301-317. Cambridge, Mass.: Belknap Press, 1987.

Jakobson, Roman. "The Metaphoric and Metonymic Poles." 1958. In *Modern Criticism and Theory: A Reader*, 3rd ed., edited by David Lodge and Nigel Wood, 165-168. Harlow: Pearson, 2008.

Jakobson, Roman. *Novejšaja Russkaja Poèzija: Nabrosok Pervyj: Viktor Xlebnikov*. Prague, 1921.

Johnson, Randal. "Editor's Introduction." In *The Field of Cultural Production*, by Pierre Bourdieu, 1-28. New York: Columbia University Press, 1993.

Johnstone, Barbara. "Language and Place." In *The Cambridge Handbook of Sociolinguistics*, edited by Rajend Mesthrie, 203-217. Cambridge: Cambridge University Press, 2011.

Kacian, Jim. "An Overview of Haiku in English." In *Haiku in English: The First Hundred Years*, edited by Jim Kacian, Philip Rowland and Allan Burns, 305-378. New York: W. W. Norton, 2013.

Ketai, R. "Affect, Mood, Emotion, and Feeling: Semantic Sonsiderations." *The American Journal of Psychiatry* 132, no. 11 (1975): 1215–1217.

Kilgarriff, A., P. Rychlý, P. Smrz, and D. Tugwell. *The Sketch Engine*. 2004. Available online at: http://sketchengine.eu.

Klein, Dan and Christopher D. Manning. "Accurate Unlexicalized Parsing." In *Proceedings of the 41st Meeting of the Association for Computational Linguistics*, 423-430. Stroudsberg: Association for Computational Linguistics, 2003.

Knowles, Gerald. *Patterns of Spoken English: An Introduction to English Phonetics*. London: Longman, 1987.

Kodama, Chiaki. "Suēden-ni-okeru haiku juyō" [the reception of haiku in Sweden]. Kita
　　Yōroppa Kenkyū 4 (2007): 33-45.

Koskinen, Kaisa. "Domestication, Foreignization and the Modulation of Affect." In
　　Domestication and Foreignization in Translation Studies, edited by Hannu
　　Kemppanen, Marja Jänis and Alexandra Belikova, 13-32. Berlin: Frank & Timme,
　　2012.

Kövecses, Zoltán. *Metaphor: A Practical Introduction*. 2nd ed. Oxford: Oxford University
　　Press, 2010.

Kowal, Sabine and Daniel O'Connell. "Pausological Research at Saint Louis University." In
　　Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler, edited
　　by Hans W. Dechert and Manfred Raupach, 61-68. The Hague: Mouton, 1980.

Kuninaka, Osamu. "Yoshihara Sachiko." In *Shin kenkyū shiryō: gendai Nihon-bungaku dai-
　　nana-kan: shi* [new research materials: contemporary Japanese literature vol. 7:
　　poetry], edited by Asai Kiyoshi, Satō Masaru, Shino Hiroshi, Torī Kunio, Matsui
　　Toshihiko, Mukawa Chūichi, and Yoshida Hirō, 430-432. Tokyo: Meiji Shoin, 2000.

Kurihara, Atsushi. "Tanikawa Shuntarō." In *Shin kenkyū shiryō: gendai Nihon-bungaku dai-
　　nana-kan: shi* [new research materials: contemporary Japanese literature vol. 7:
　　poetry], edited by Asai Kiyoshi, Satō Masaru, Shino Hiroshi, Torī Kunio, Matsui
　　Toshihiko, Mukawa Chūichi, and Yoshida Hirō, 367-375. Tokyo: Meiji Shoin, 2000.

LaFleur, William. "The Expanse and the Limits of a New Anthology," review of *From the
　　Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, *Monumenta
　　Nipponica* 38, no. 2 (1983): 191-205.

Larkin, Willard and David Burns. "Sentence Comprehension and Memory for Embedded
　　Structure." *Memory & Cognition* 5, no. 1 (1977): 17-22.

Leach, Laura and Arthur G. Samuel. "Lexical Configuration and Lexical Engagement: When

Adults Learn New Words." *Cognitive Psychology* 55, no. 4 (2007): 306–353.

Lecercle, Jean-Jacques. *The Violence of Language*. London: Routledge, 1990.

Lee, David. "BNC Web Indexer." 2013. Available online at: http://ucrel.lancs.ac.uk/
bncindex/.

Leech, Geoffrey. "A Brief Users' Guide to the Grammatical Tagging of the British National
Corpus." 1995. http://www.natcorp.ox.ac.uk/docs/gramtag.html. Accessed 27 May
2019.

Leech, Geoffrey, Paul Rayson and Andrew Wilson. *Word Frequencies in Written and Spoken
English*. Harlow: Longman, 2001.

Leitch, Vincent B., William E. Cain, Laurie A. Finke, Barbara E. Johnson, John McGowan,
T. Denean Sharpley-Whiting and Jeffrey J. Williams, eds. *The Norton Anthology of
Theory and Criticism*. 2nd ed. New York: W. W. Norton, 2010.

Lemarié, Julie, Robert F. Lorch Jr. and Marie-Paule Péry-Woodley. "Understanding How
Headings Influence Text Processing." *Discours* 10 (2012): 1-22.

Lenneberg, E. H. *Biological Foundations of Language*. New York: Wiley, 1967.

Lento, Takako U., trans. "Alone in Two Billion Light Years." In *The Art of Being Alone:
Poems 1952-2009*, by Shuntarō Tanikawa, 47. New York: Cornell East Asia Series,
2011.

Lewis, Philip. "The Measure of Translation Effects." In *Difference in Translation*, edited by
Joseph F. Graham, 31-62. Ithaca: Cornell University Press, 1985.

Louw, Bill. "Contextual Prosodic Theory: Bringing Semantic Prosodies to Life." In *Words in
Context*, edited by Chris Heffer and Helen Sauntson, 48-94. Birmingham: University
of Birmingham, 2000.

Louw, Bill. "Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of
Semantic Prosodies." In *Text and Technology*, edited by Mona Baker, Gill Francis

and Elena Tognini-Bonelli, 157-175. Philadelphia: John Benjamins, 1993.

Louw, Bill and Marija Milojkovic. *Corpus Stylistics as Contextual Prosodic Theory and Subtext*. Amsterdam: John Benjamins, 2016.

Ludden, David. *The Psychology of Language: An Integrated Approach*. London: SAGE, 2016.

Makel, Matthew C., Jonathan A. Plucker and Boyd Hegarty. "Replications in Psychology Research: How Often Do They Really Occur?" *Perspectives on Psychological Science* 7, no. 6 (2012): 537-542.

Marling, William. *Gatekeepers: The Emergence of World Literature and the 1960s*. New York: Oxford University Press, 2016.

Masaoka Shiki. *Shiki zenshū daiyonkan: hairon haiwa I* [complete works of Shiki, vol. 4: theory and discussion of haiku part 1]. Tokyo: Kōdansha, 1979.

Matsui, Takako. "Masaoka Shiki-no Bashō rikai" [Masaoka Shiki's understanding of Bashō]. *Renga Haikai Kenkyū* 90 (1996): 28-37.

Matsumoto, David, ed. *The Cambridge Dictionary of Psychology*. Cambridge: Cambridge University Press, 2009.

Mazurek Melnyk, Bernadette, Dianne Morrison-Beedy and Shirley M. Moore. "Nuts and Bolts of Designing Intervention Studies." In *Intervention Research: Designing, Conducting, Analyzing, and Funding*, edited by Bernadette Mazurek Melnyk and Dianne Morrison-Beedy, 37-64. New York: Springer, 2012.

McDonald, John H. "*G*-test of goodness-of-fit." In *Handbook of Biological Statistics*, 3rd ed. Available online at: http://www.biostathandbook.com/gtestgof.html. Baltimore: Sparky House Publishing, 2014.

McKillup, Steve. *Statistics Explained: An Introductory Guide for Life Scientists*. 2nd ed. Cambridge: Cambridge University Press, 2012.

McLaughlin, Jacqueline. "A Gentle Reminder that a Hypothesis Is Never Proven Correct, nor Is a Theory Ever Proven to Be True." *Journal of College Science Teaching* 36, no. 1 (2006): 60-62.

Mellinger, Christopher D. and Thomas A. Hanson. *Quantitative Research Methods in Translation and Interpreting Studies*. London: Routledge, 2017.

Miall, David S. *Literary Reading: Empirical & Theoretical Studies*. New York: Peter Lang, 2006.

Miall, David S. and Don Kuiken. "Foregrounding, Defamiliarization and Affect: Response to Literary Stories." *Poetics* 22, no. 5 (1994): 389-407.

Miller, Brian and Taeko Kudo, trans. "Instant," by Yoshihara Sachiko. In *Three Poets: Emori Kunitomo, Konagaya Kiyomi, Yoshihara Sachiko*, 57. Tokyo: BT Books, 1987.

Miller, Roy Andrew. *The Japanese Language*. Chicago: University of Chicago Press, 1967.

Miller, Steve. *Experimental Design and Statistics*. 2nd ed. London: Routledge, 1984.

Milroy, Leslie and Matthew Gordon. *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell, 2003.

Miner, Earl. Review of *From the Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, *The Journal of Japanese Studies* 8, no. 2 (1982): 388-391.

Minsky, Marvin. "A Framework for Representing Knowledge." In *The Psychology of Computer Vision*, edited by Patrick Henry Winston, 211-277. New York: McGraw-Hill, 1975.

Miranda, Claudia. "'*Dove*' is the Dove?" In *Reading Eco: An Anthology*, edited by Rocco Capozzi, 362-386. Bloomington: Indiana University Press, 1997.

Morton, Leith, trans. "CLAMS," by Rin Ishigaki. *Poetry International* 2005. Available online at: https://www.poetryinternational.org/pi/poem/3777/auto/0/0/Rin-Ishigaki/CLAMS/en/tile.

Mountain, Marlene. *The Old Tin Roof*. Hampton, TN: Marlene Mountain, 1976. Available online at: https://www.thehaikufoundation.org/omeka/items/show/60.

Mukařovský, Jan. "Standard Language and Poetic Language." 1932. In *A Prague School Reader on Esthetics, Literary Structure, and Style*, translated by Paul L. Garvin, 17-30. Washington, D.C.: Georgetown University Press, 1964.

Muller, Johan. "Social Life in Disciplines." *In Knowledge and Identity: Concepts and Applications in Bernstein's Sociology*, edited by Gabrielle Ivinson, Brian Davies and John Fitz, 39-53. London: Routledge 2011.

Munday, Jeremy. *Introducing Translation Studies: Theories and Applications*. 2nd ed. London: Routledge, 2008.

Muth, Claudia and Claus-Christian Carbon. "The Aesthetic Aha: On the Pleasure of Having Insights into Gestalt." *Acta Psychologica* 144, no. 1 (2013): 25-30.

Myskja, Kjetil. "Foreignization and Resistance: Lawrence Venuti and His Critics." *Nordic Journal of English Studies* 12, no. 2 (2013): 1-23.

Naff, William E. "The Heritage of Courtly Love in Japanese Society." *Comparative Civilizations Review* no. 1 (1979): 42-63.

Nakayasu, Sawako with Chika Sagawa. *Mouth: Eats Color—Sagawa Chika Translations, Anti-Translations & Originals*. Tokyo: Rogue Factorial, 2011.

Nørgaard, Nina, Rocío Montoro, and Beatrix Busse. *Key Terms in Stylistics*. London: Continuum, 2010.

"Obligative Mood." In *SIL Glossary of Linguistic Terms*, edited by Eugene E. Loos, Susan Anderson, Dwight H. Day, Jr., Paul C. Jordan and J. Douglas Wingate. SIL International, 2003. https://glossary.sil.org/term/obligative-mood. Accessed 25 July 2018.

Ochiai, Keiko. "Kotoba-de, tōge-wo koeru" [overcoming the worst through words]. In

Ishigaki Rin shishū [selected poems of Ishigaki Rin], 243-246. Tokyo: Haruki Bunko, 1998.

Ogunnaike, Babatunde A. *Random Phenomena: Fundamentals of Probability and Statistics for Engineers*. Boca Raton: CRC Press, 2010.

Ōoka Makoto. "Yoshihara Sachiko-no shi" [the poetry of Yoshihara Sachiko]. In *Zoku Yoshihara Sachiko shishū* [selected poems of Yoshihara Sachiko: continued], 142-149. Tokyo: Shichōsha, 2003.

O'Toole, Michael. "Russian Literary Theory: From the Formalists to Lotman." In *The Routledge Companion to Russian Literature*, edited by Neil Cornwell, 163-173. London: Routledge, 2001.

Pan, Joe and Martin Rock, trans. "From *River / Horse / Haiku*," by Tsubouchi Nenten. *Asymptote* January 20. https://www.asymptotejournal.com/special-feature/tsubouchi-nenten-river-horse-haiku/.

Pelowski, Matthew, Patrick S. Markey, Michael Forster, Gernot Gerger and Helmut Leder. "Move Me, Astonish Me. . . Delight My Eyes and Brain: The Vienna Integrated Model of Top-Down and Bottom-Up Processes in Art Perception (VIMAP) and Corresponding Affective, Evaluative, and Neurophysiological Correlates." *Physics of Life Reviews* 21 (2017): 80-125.

Peplow, David and Ronald Carter. "Stylistics and Real Readers." In *The Routledge Handbook of Stylistics*, edited by Michael Burke, 440-454. London: Routledge, 2014.

Perea, Manuel, Mariko Nakayama and Stephen J. Lupker. "Alternating-Script Priming in Japanese: Are Katakana and Hiragana Characters Interchangeable?" *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43, no. 7 (2017): 1140 – 1146.

Pett, Marjorie A. *Nonparametric Statistics for Health Care Research*. Los Angeles: SAGE,

1997.

Philip, Gill. *Colouring Meaning: Collocation and Connotation in Figurative Language*. Amsterdam: John Benjamins, 2011.

Phillips, Martin. *Aspects of Text Structure: An Investigation of the Lexical Organisation of Text*. Amsterdam: North-Holland, 1985.

Piantadosi, Steven T., Harry Tily and Edward Gibson. "Word Lengths Are Optimized for Efficient Communication." *Proceedings of the National Academy of Sciences of the United States of America* 108, no. 9 (2011): 3526–3529.

Pisanty, Valentina. "The Sememe as a Virtual Test." In *Umberto Eco in His Own Words*, edited by Torkild Thellefsen and Bent Sørensen, 143-148. Boston: De Gruyter, 2017.

Polley Sanchez, Rebecca, Elizabeth Pugzles Lorch and Robert F. Lorch, Jr. "Effects of Headings on Text Processing Strategies." *Contemporary Educational Psychology* 26, no. 3 (2001): 418–428.

Pomorska, Krystyna. "Introduction." In *Language in Literature*, by Roman Jakobson, 1-11. Cambridge, Mass.: Belknap Press, 1987.

Portner, Paul. "Vocatives, Topics, and Imperatives." Lecture, IMS Workshop on Information Structure, Bad Teinach, Germany, 16 July 2004. Lecture handout available online at https://faculty.georgetown.edu/portnerp/Papers/VocativesHandout.pdf. Accessed 25 July 2018.

Pratt, Mary Louise. *Toward a Speech Act Theory of Literary Discourse*. Bloomington: Indiana University Press, 1977.

Privitera, Gregory J. *Statistics for the Behavioral Sciences*. Los Angeles: SAGE, 2012.

Pym, Anthony. "Venuti's Visibility," review of *The Translator's Invisibility: A History of Translation*, by Lawrence Venuti, *Target* 8, no. 1 (1996): 165-177.

Quartermain, Peter. *Stubborn Poetries: Poetic Facticity and the Avant-Garde*. Tuscaloosa:

The University of Alabama Press, 2013.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. *A Comprehensive Grammar of the English Language*. London: Longman, 1985.

R Core Development Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2018.

Radvansky, Gabriel A. and Katinka Dijkstra. "Aging and Situation Model Processing." *Psychonomic Bulletin & Review* 14, no. 6 (2007): 1027-1042.

Radvansky, Gabriel A., Rolf A. Zwaan, Jacqueline M. Curiel and David E. Copeland. "Situation Models and Aging." *Psychology & Aging* 16, no. 1 (2001): 145-160.

Ramazani, Jahan. *Poetry in a Global Age*. Chicago: The University of Chicago Press, 2020.

Rayner, Keith, Timothy J. Slattery and Nathalie N. Bélanger. "Eye Movements, the Perceptual Span, and Reading Speed." *Psychonomic Bulletin & Review* 17, no. 6 (2010): 834–839.

Rayson, Paul. "Log-Likelihood and Effect Size Calculator." 2016. Available online at: http://ucrel.lancs.ac.uk/llwizard.html.

Rayson, Paul. *Wmatrix*. 2009. Available online at: http://ucrel.lancs.ac.uk/wmatrix/.

Reber, Rolf. "Processing Fluency, Aesthetic Pleasure and Culturally Shared Taste." In *Aesthetic Science: Connecting Minds, Brains, and Experience*, edited by Arthur P. Shimamura and Stephen E. Palmer, 223-249. New York: Oxford University Press, 2012.

Reber, Rolf, Norbert Schwarz and Piotr Winkielman. "Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience?" *Personality and Social Psychology Review* 8 (2004): 364–82.

Reichhold, Jane and Machiko Kobayashi, trans. *Midaregami* no. 145, by Yosano Akiko. In *A Girl with Tangled Hair*, by Yosano Akiko, 99. Gualala, CA: AHA Books, 2014.

Rexroth, Kenneth and Ikuko Atsumi, trans. "Shellfish" by Rin Ishigaki. In *Women Poets of Japan*, translated by Kenneth Rexroth and Ikuko Atsumi. New York: New Directions, 1977.

Reynolds, Matthew. "Introduction." In *Prismatic Translation*, edited by Matthew Reynolds, 1-18. Cambridge: Legenda, 2019.

Reynolds, Matthew. *Translation: A Very Brief Introduction*. Oxford: Oxford University Press, 2016.

Reynolds, Matthew, Sowon S. Park and Kate Clanchy. "Prismatic Translation." In *Creative Multilingualism: A Manifesto*, edited by Katrin Kohl, Rajinder Dudrah, Andrew Gosler, Suzanne Graham, Martin Maiden, Wen-chin Ouyang and Matthew Reynolds, 131-149. Cambridge: Open Book Publishers, 2020.

Richards, I. A. *Practical Criticism: A Study of Literary Judgment*. London: Kegan Paul, Trench, Trubner & Co., 1929.

Rimer, J. Thomas and Van C. Gessel, eds. *The Columbia Anthology of Modern Japanese Literature*, vol. 2. New York: Columbia University Press, 2005.

Roach, Peter. "British English: Received Pronunciation." *Journal of the International Phonetic Association* 34, no. 2 (2004): 239-245.

Roach, Peter. *English Phonetics and Phonology: A Practical Course*. 4th ed. Cambridge: Cambridge University Press, 2009.

Robertson, Wes. "Orthography, Foreigners, and Fluency: Indexicality and Script Selection in Japanese Manga." *Japanese Studies* 35, no. 2 (2015): 205-222.

Robey, David. "Introduction: Interpretation and Uncertainty." In *Illuminating Eco: On the Boundaries of Interpretation*, edited by Charlotte Ross and Rochelle Sibley, 1-12. Oxon: Routledge, 2004.

Robinson, Douglas. *Estrangement and the Somatics of Literature: Tolstoy, Shklovsky, Brecht*.

Baltimore: The Johns Hopkins University Press, 2008.

Robinson, Douglas. *Schleiermacher's Icoses: Social Ecologies of the Different Methods of Translating*. Bucharest: Zeta Books, 2013.

Robinson, Douglas. *Translation and the Problem of Sway.* Amsterdam: John Benjamins, 2011.

Robinson, Douglas. *Translationality: Essays in the Translational-Medical Humanities*. London: Routledge, 2017.

Rosenberg, M. J. and C.I. Hovland, C. I. "Cognitive, Affective, and Behavioral Components of Attitudes." In *Attitude Organization and Change*, edited by M. Rosenberg , C. Hovland , W. McGuire , R. Abelson , & J. Brehm, 1-14. New Haven: Yale University Press, 1960.

Royal Society of Literature. 2017a. "Literature in Britain Today." London: The Royal Society of Literature, 2017.

Royal Society of Literature. 2017b. "Literature in Britain Today" Data Tables. Available online at: https://www.ipsos.com/sites/default/files/migrations/en-uk/files/Assets/Docs/Polls/omnibus-rsl-tables-2017.pdf.

Saeed, John I. *Semantics*. 2nd ed. Oxford: Blackwell, 2003.

Saldanha, Gabriela. 2011a. "Emphatic Italics in English Translations: Stylistic Failure or Motivated Stylistic Resources." *Meta* 56, no. 2 (2011): 424-442.

Saldanha, Gabriela. 2011b. "Translator Style: Methodological Considerations." *The Translator* 17-1 (2011): 25-50.

Salkind, Neil J. "Convenience Sampling." In *Encyclopedia of Research Design*, edited by Neil J. Salkind, 254. Los Angeles: SAGE, 2010.

San, Debra. "Hiatus of Subject and Verb in Poetic Language." *Style* 39, no. 2 (2005): 137-152.

Sanford, Anthony J. and Patrick Sturt. "Depth of Processing in Language Comprehension: Not Noticing the Evidence." *Trends in Cognitive Sciences* 6, no. 9 (2002): 382-386.

Saris, Willem E. and Irmtraud N. Gallhofer. *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, NJ: Wiley, 2007.

Sato, Hiroaki, trans. 1981a. "Clams," by Rin Ishigaki. In *From the Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, 574. Seattle: University of Washington Press, 1981.

Sato, Hiroaki, trans. 1981b. *Midaregami* no. 145, by Yosano Akiko. In *From the Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, 433. Seattle: University of Washington Press, 1981.

Schank, Roger C. and Roger P. Abelson. *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1977.

Scholes, Robert. "Who Cares About the Text?" *NOVEL: A Forum on Fiction* 17, No. 2 (1984): 171-180.

Scollon, R. and S. Scollon. "The Literate Two-Year-Old: The Fictionalization of Self." *Working Papers in Sociolinguistics*. Austin, TX: Southwest Regional Laboratory.

Scott, Mike. "Focusing on the Text and Its Key Words." In *Rethinking Language Pedagogy from a Corpus Perspective: Papers from the Third International Conference on Teaching and Language Corpora*, edited by Lou Burnard and Tony McEnery, 103-122. Frankfurt am Main: Peter Lang, 2000.

Scott, Mike. *Oxford WordSmith Tools 4.0*. Oxford: Oxford University Press, 2007.

Scott, Mike. "PC Analysis of Key Words – And Key Key Words." *System* 25, no. 2 (1997): 233-245.

Seeley, Christopher. *A History of Writing in Japan.* Honolulu: University of Hawai'i Press, 1991.

Shamma, Tarek. "The Exotic Dimension of Foreignizing Strategies: Burton's Translation of the *Arabian Nights*." *The Translator* 11, no. 1 (2005): 51-67.

Shang, Miles. 2011. "Syntax Tree Generator." http://mshang.ca/syntree/. Accessed 27 May 2019.

Sharpe, R. A. "Solid Joys or Fading Pleasures." In *Pleasure, Preference and Value*, edited by Eva Shaper, 86-98. Cambridge: Cambridge University Press, 1983.

Sher, Benjamin, trans. *Theory of Prose*, by Viktor Shklovsky. Elmwood Park, IL: Dalkey Archive Press, 1990.

Shimizu, Yoshinori. "Kokugo nyūshi mondai hisshō-hō." In *Kokugo nyūshi mondai hisshō-hō*, 31-58. Tokyo: Kodansha, 1990.

Shklovsky, Victor. "Art as Technique." 1917. In *Russian Formalist Criticism: Four Essays*, translated and with an introduction by Lee T. Lemon and Marion J. Reis, 3-24. Lincoln: University of Nebraska Press, 1965.

Shklovsky, Viktor. "Art as Device." 1917/1919. In *Viktor Shklovsky: A Reader*, edited and translated by Alexandra Berlina, 73-96. New York: Bloomsbury, 2017.

Shklovsky, Viktor. *Literatura i kinematograf* [literature and cinematography]. Berlin: Russkoe Universal'noe Izdatel'stvo, 1923.

Shklovsky, Viktor. *Literature and Cinematography*. 1923. Translated by Irina Masinovsky. Champaign : Dalkey Archive Press, 2008.

Shklovsky, Viktor. "Literature Beyond 'Plot.'" 1921/1925. In *Viktor Shklovsky: A Reader*, edited and translated by Alexandra Berlina, 97-104. New York: Bloomsbury, 2017.

Shklovsky, Viktor. "Resurrecting the Word." 1914. In *Viktor Shklovsky: A Reader*, edited and translated by Alexandra Berlina, 63-72. New York: Bloomsbury, 2017.

Shklovsky, Viktor. *Rozanov*. Petersburg: OPOJAZ, 1921.

Shklovsky, Viktor. Excerpts from *A Sentimental Journey*. 1923. In In *Viktor Shklovsky: A*

Reader, edited and translated by Alexandra Berlina, 135-152. New York: Bloomsbury, 2017.

Shklovsky, Viktor. *Sentimental'noye puteshestviye* [a sentimental journey]. Berlin: Gelikon, 1923.

Shreve, Gregory M. "Text Linguistics, Translating, and Interpreting." In *The Routledge Handbook of Translation Studies and Linguistics*, edited by Kirsten Malmkjær, 165-178. London Routledge, 2018.

Simpson, Paul. *Stylistics: A Resource Book for Students*. London: Routledge, 2004.

Sinclair, John. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press, 1991.

Sinclair, John. *Looking Up*. London: Collins, 1987.

Smith, Barbara Herrnstein. "On the Margins of Discourse." *Critical Inquiry* 1, no. 4 (1975): 769-798.

Smith, Barbara Herrnstein. *On the Margins of Discourse: The Relation of Literature to Language*. Chicago: University of Chicago Press, 1979.

Smith, M. Cecil. "The Reading Abilities and Practices of Older Adults." *Educational Gerontology* 19, no. 5 (1993): 417-432.

Spencer, Hannah. "Semantic Prosody in Literary Analysis: A Corpus-Based Stylistic Study of H. P. Lovecraft's Stories." Master's thesis, University of Huddersfield, 2011.

Spencer, Herbert. *The Philosophy of Style*. 1852. New York: D. Appleton and Company, 1884.

Stankiewicz, Edward. "Linguistics and the Study of Poetic Language." In *Style in Language*, edited by Thomas A. Seobok, 69-81. Cambridge, MA: The M.I.T. Press, 1960.

Steiner, Peter. *Russian Formalism: A Metapoetics*. Ithaca: Cornell University Press, 1984.

Stewart, Dominic. *Semantic Prosody: A Critical Evaluation*. London: Routledge, 2010.

Striedter, Jurij. *Literary Structure, Evolution, and Value: Russian Formalism and Czech*

Structuralism Reconsidered. Cambridge, MA: Harvard University Press, 1989.

Strohl, Matthew. "Horror and Hedonic Ambivalence." *The Journal of Aesthetics and Art Criticism* 70, no. 2 (2012): 203-212.

Suzuki, Sadami. "Higashi-Ajia kingendai-no gainen henseishi kenkyū-no genzai" [present state of research on the history of modern and contemporary concepts in East Asia]. In *Higashi-Ajia-ni-okeru gakugei-shi-no sōgō-teki kenkyū-no keizoku-teki hatten-no tame ni*, 7-20. Kyoto: International Research Center for Japanese Studies, 2013.

Suzuki, Sadami. *Nihon-no 'bungaku' gainen* [the concept of 'literature' in Japan]. Tokyo: Sakuhinsha, 1998.

Taboada, María Teresa. *Building Coherence and Cohesion*. Amsterdam: John Benjamins, 2004.

Tanikawa Shuntarō. "Nijūoku kōnen-no kodoku" [loneliness of two billion light years]. 1952. In *Gendai shishū* [anthology of contemporary poetry], 401. *Gendai Nihon bungaku zenshū*, vol. 89. Tokyo: Chikuma Shobō, 1967.

Taylor-Jones, Kate. "Japan Now North Festival Report." *East Asian Journal of Popular Culture* 5, no. 2 (2019): 203-206.

Teddiman, Laura. "Subject Ellipsis by Text Type: An Investigation Using ICE-GB." In *Corpus-Based Studies in Language Use, Language Learning, and Language Documentation*, edited by John Newman, Harald Baayan and Sally Rice, 71-90. Amsterdam: Rodopi, 2011.

Ten Hacken, Pius. *Chomskyan Linguistics and Its Competitors.* London: Equinox, 2007.

Thwaite, Anthony. "Haiku, Tanka, Hokku, Senryu, Choka," review of *From the Country of Eight Islands*, edited by Hiroaki Sato and Burton Watson, *New York Times*, February 15, 1981, Section 7, Page 13.

toPhonetics [sic]. Online IPA transcription tool. Available at: https://tophonetics.com/.

Toth, Hayley Georgia. "Contemporary Postcolonial Literature, Reader-Response, and Reception Studies." Ph.D. diss, University of Leeds, 2019.

Townsend, David J., David Ottaviano and Thomas G. Bever. "Immediate Memory for Words from Main and Subordinate Clauses at Different Age Levels." *Journal of Psycholinguistic Research* 8, no. 1 (1979): 83-101.

Trask, R. L. *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge, 1992.

Trudgill, Peter. *Sociolinguistic Variation and Change.* Washington, D.C.: Georgetown University Press, 2002.

Tsujimura, Natsuko. *An Introduction to Japanese Linguistics*. 3rd ed. Malden, MA: Wiley, 2014.

Tsur, Reuven. "A Conversation with Reuven Tsur on Intonation and Prosody." By Natalie Gerber and David Nowell Smith. *Thinking Verse* 5 (2015): 172-213.

Tymoczko, Maria. "Translation and Political Engagement." *The Translator* 6, no.1 (2000): 23-47.

Tyng, Chai M., Hafeez U. Amin, Mohamad N. M. Saad and Aamir S. Malik. "The Influences of Emotion on Learning and Memory." *Frontiers in Psychology* 8 (2017): 1-22.

UCREL (University Centre for Computer Corpus Research on Language). 1996. "A Post-Editor's Guide to Claws7 Tagging." http://www.natcorp.ox.ac.uk/docs/claws7.html. Accessed 27 May 2019.

UCREL (University Centre for Computer Corpus Research on Language). No date. "UCREL CLAWS7 Tagset." http://ucrel.lancs.ac.uk/claws7tags.html. Accessed 27 May, 2019.

Van Peer, Willie. *Stylistics and Psychology: Investigations of Foregrounding*. London: Croom Helm, 1986.

Van Peer, Willie, Frank Hakemulder and Sonia Zyngier. *Scientific Methods for the Humanities*. Amsterdam: John Benjamins, 2012.

Van Peer, Willy [sic]. *The Stylistic Theory of Foregrounding: A Theoretical and Empirical Investigation*. PhD diss., University of Lancaster, 1980.

Vaughan, Liwen. *Statistical Methods for the Information Professional*. Medford, NJ: Information Today, Inc., 2001.

Venuti, Lawrence. *Contra Instrumentalism: A Translation Polemic*. Lincoln: University of Nebraska Press, 2019.

Venuti, Lawrence. *The Scandals of Translation: Towards an Ethics of Difference*. London: Routledge, 1998.

Venuti, Lawrence. *Translation Changes Everything: Theory and Practice*. London: Routledge, 2013.

Venuti, Lawrence. "The Translator's Invisibility." *Criticism* 28, no. 2 (1986): 179-212.

Venuti, Lawrence. *The Translator's Invisibility: A History of Translation*. 1st ed. London: Routledge, 1995.

Venuti, Lawrence. *The Translator's Invisibility: A History of Translation*. 2nd ed. London: Routledge, 2008.

Venuti, Lawrence. *The Translator's Invisibility: A History of Translation*. 3rd ed. New York: Routledge, 2018.

Vincent, J. Keith. Review of *On Haiku*, by Hiroaki Sato, *Translation Review* 106 (2020): 108-119.

Wales, Katie. *A Dictionary of Stylistics*. 3rd ed. London: Routledge, 2011.

Waley, Arthur. "Japanese." In *Eos: An Enquiry into the Theme of Lovers' Meetings and Partings at Dawn in Poetry*, edited by Arthur T. Hatto, 114-125. London: Mouton, 1965.

Wallace, Walter L. *The Logic of Science in Sociology*. New York: De Gruyter, 1971.

Wallace, Walter L. "'Why Sociology Doesn't Make Progress.'" *Sociological Forum* 10, no. 2

(1995): 313-318.

Wang, J., L. Li, S. Li, F. Xie, S.P. Liversedge and K.B. Paterson. "Effects of Aging and Text-Stimulus Quality on the Word-Frequency Effect During Chinese Reading." *Psychology and Aging* 33, no. 4 (2018): 693–712.

Wellek, René. "The New Criticism: Pro and Contra." *Critical Inquiry* 4, no. 4 (1978): 611-624.

Wells, J. C. *English Intonation: An Introduction.* Cambridge: Cambridge University Press, 2006.

West, David W. "Practical Criticism: I.A. Richards' Experiment in Interpretation." *Changing English* 9, no. 2 (2002): 207-213.

Widdowson, Peter. *Literature*. London: Routledge, 1999.

Wiley, Jennifer and Keith Rayner. "Effects of Titles on the Processing of Text and Lexically Ambiguous Words:Evidence from Eye Movements." *Memory & Cognition* 28, no. 6 (2000): 1011-1021.

Williams, Rihana and Robin Morris. "Eye Movements, Word Familiarity, and Vocabulary Acquisition." *European Journal of Cognitive Psychology* 16, nos. 1/2 (2004): 312–339.

Wimsatt, W.K., Jr., and M.C. Beardsley. "The Affective Fallacy." *The Sewanee Review* 57, no. 1 (1949): 31-55.

Wimsatt, W.K., Jr., and M.C. Beardsley. "The Intentional Fallacy." *The Sewanee Review* 54, no. 3 (1946): 468-488.

Wochna, Kacey L. and Barbara J. Juhasz. "Context Length and Reading Novel Words: An Eye-Movement Investigation." *British Journal of Psychology* 104, no. 3(2013): 347–363.

Wolf, Maryanne. *Proust and the Squid: The Story and Science of the Reading Brain*. London:

Icon Books, 2008.

Yablo, Stephen. *Aboutness*. Princeton: Princeton University Press, 2014.

Yablo, Stephen. "Non-Catastrophic Presupposition Failure." In *Content and Modality*, edited by Judith Thomson and Alex Byrne, 164-190. Oxford: Clarendon Press, 2006.

Yamamoto, Michiko. "Yoshihara Sachiko-ron" [discourse on Yoshihara Sachiko]. In *Yoshihara Sachiko shishū* [selected poems of Yoshihara Sachiko], 157-160. Tokyo: Shichōsha, 1976.

Yao, Xinyue and Peter Collins. "The Present Perfect in World Englishes." *World Englishes* 31, no. 3 (2012): 386-403.

Yoo, Eun-Jung. "Floating Quantifiers and Lexical Specification of Quantifier Retrieval." *In Proceedings of the 15th Pacific Asia Conference on Language, Information and Computation*, edited by Benjamin K. T'sou, Olivia O.Y. Kwong and Tom B.Y. Lai, 211-222. Hong Kong: City University of Hong Kong, 2001.

Yosano, Akiko. *Midaregami* [tangled hair]. 1901. Tokyo: Shichō Bunko, 2000.

Yoshida, Hirō. "Kaisetsu" [commentary]. In *Kindai dokusha-no seiritsu* [the creation of modern readers], by Ai Maeda, 465-480. Tokyo: Chikuma Shobō, 1989.

Yoshihara, Sachiko. "Shunkan" [instant]. 1964. In *Yoshihara Sachiko zenshi I* [collected poems of Yoshihara Sachiko volume 1], 131-133. Tokyo: Shichōsha, 1981.

Yoshihara, Sachiko. "Tōkai-ni tsuite" [on concealing oneself]. 1975. In *Yoshihara Sachiko*, *Gendai-no shijin* vol. 12, edited by Ōoka Makoto and Tanikawa Shuntarō, 189-197. Tokyo: Chūō Kōronsha, 1983.

Yule, George. *The Study of Language*. 6th ed. Cambridge: Cambridge University Press, 2016.

Zanuttini, Rafaella. "Encoding the Addressee in the Syntax: Evidence from English Imperative Subjects." *Natural Language & Linguistic Theory* 26, no. 1 (2008): 185-218.

Zyngier, Sonia, Willie Van Peer, and Jèmeljan Hakemulder. "Complexity and

    Foregrounding: In the Eye of the Beholder?" *Poetics Today* 28, no. 4 (2007): 653-

    682.