Machine translation, ethics and the literary translator’s voice

*Dorothy Kenny & **Marion Winters
*Dublin City University
**Heriot-Watt University

Abstract
Recent work in translation studies has established the literary translator’s voice as an ethical concern, but there has been little empirical research so far into how the translator’s voice is affected in workflows involving machine translation. In this article, we investigate how the use of neural machine translation influences the textual voice (Alvstad et al. 2017) of renowned translator from English into German, Hans-Christian Oeser. Based on an experiment in which Oeser post-edits an excerpt from a novel he had previously translated, we show how his textual voice is somewhat diminished in his post-edited work compared to its stronger manifestation in his translation work. At the same time Oeser’s contextual voice (ibid.) remains strong in his comments on the text he produces in post-editing mode. The article is offered as a methodological intervention and represents an initial attempt to design studies in literary machine translation that put the focus on human translators, allowing their voices to be heard more clearly than has previously been the case.

Keywords
literary translation, translator’s voice, textual voice, contextual voice, translator style, neural machine translation

Introduction
This article starts from the premise that there is such a thing as a translator’s ‘voice’, understood, after Hermans (1996, 27), as “an index of the Translator’s discursive presence” in a text. Although we may be conditioned to expect direct access to an original author’s (or narrator’s) voice – and that voice only – in translations, there are occasions on which the univocal frame is broken, and the reader is jolted into an awareness of the text’s plurivocal nature (ibid. 33). A translator’s note, for example, immediately draws the reader’s attention to the fact that the original narrator is not the only enunciating subject ‘speaking’ in a text. There are other, more subtle, ways in
which a translator’s presence may suddenly become discernible, but to have been flushed out, it must have always been there in the first place.

Hermans’ (1996) uncovering of the translator’s voice relies on the identification of explicit paratextual interventions or dissonances within the target text. The translator’s discursive presence can also be argued on theoretical, and especially narratological, grounds (see Mossop 1983; Schiavi 1996), or induced from translations completed by particular translators (Folkart 1991; Pekkanen 2010). Scholars working within corpus-based translation studies (e.g. Baker 2000; Munday 2008; Saldanha 2011; Winters 2004, 2007, 2009) adopt the latter approach, but use large digital collections of translations to provide empirical evidence that individual translators leave a kind of “thumbprint” on their translations, in the form of consistent stylistic preferences that are manifest in target texts. In more recent years, translation scholarship concerned with voice has expanded to accommodate the large number of other protagonists whose voices can be heard in and around translations. They include not just authors and translators, but publishers, editors, copy editors, and critics, to name just a few. This line of research has been championed by the group of scholars involved in the *Voices of Translation* project, and especially Cecilia Alvstad, Annjo Greenall, Hanne Jansen and Kristiina Taivalkoski-Shilov (Alvstad et al. 2017; Greenall et al. 2019).

Despite the dynamism of the field, there has, to our knowledge, been only one study so far that specifically addresses the concept of voice in the context of machine translation: Taivalkoski-Shilov (2019; and see below) opens a rich vein by broaching ethical issues that emerge in the machine(-assisted) translation of literary texts, of which voice is but one. In this article we attempt to build on her work, by providing an empirical analysis of one translator’s voice, as manifested in his post-editing work. We adopt a quasi-experimental design in which the translator is asked to post-edit a literary translation performed by a machine translation system, and we compare his post-edited version to both the machine translation output, and to a previous translation of the same excerpt that he produced without any machine translation aid. We are interested in whether we can detect traces of his voice in the edits he makes to the machine-translated text, but we also wish to scrutinize the opportunities he takes to comment on the process, to have a say, in more general terms. The article is structured as follows: we first review the literature on voice, ethics and (machine) translation, before describing our research design. We then present the results of our
study, and our tentative conclusions. Our work is offered as a methodological intervention, representing as it does an initial attempt to design a study in literary machine translation that gives voice to the human translator.

**Voice and its ethical import in translation**

The recognition of voice in translation, and especially the translator’s voice, is an inherently ethical move. To acknowledge the translator’s voice is to attribute to the translator agency and responsibility for particular discursive acts (Hermans 1996, 33). By recognizing plurivocality in translation we make multiple actors accountable (Alvstad et al. 2017, 4) and admit the potential for conflict (Greenall et al. 2019, 639-640). Ethical issues pertaining to voice are also to the fore in Taivalkoski-Shilov’s (2019) discussion of machine(-assisted) translation of literary texts. Here the discussion starts from the observation that while all discourse “is multi-voiced and refers to earlier discourses,” (ibid. 695; following Bakhtin 1981), multivocality and heteroglossia are particularly characteristic of literary texts, where they serve important functions. Voice is omnipresent, complex and partially constructed “between the lines” (ibid. 696), all of which makes it particularly vulnerable to mishandling when literary texts are translated by machines. Of particular concern – and extrapolating somewhat from Taivalkoski-Shilov (2019) – is the potential homogenization of different authors’ styles in contexts where a single machine translation engine is used to translate the work of multiple authors, or, conversely, the heterogenization of a single author’s style in contexts where the data used to train a machine translation engine come from multiple different sources.

Although Taivalkoski-Shilov draws attention to the translator’s voice in her article, reminding us that in a translated literary text it is the translator’s voice that becomes the narrating instance (ibid. 696), she does not expand on the ways in which this voice might be compromised in workflows involving machine translation. She does, however, elaborate on the machine translation’s voice, noting that “each MT system also has a voice” (ibid. 697) but that the machine’s voice is normally received as aesthetically and morally unacceptable “noise”. Against this background, the interventions of those writers who make an aesthetic virtue of such noise stand out. Taivalkoski-Shilov focuses on Lee’s (2011) treatment of the Taiwanese poet Hsia Yü in this regard, but a number of other writers have prized machine translation for its idiosyncrasies, its “non-linearity, unpredictability, and lack of common sense” in
Taivalkoski-Shilov’s (ibid. 698) words. Kenny (2020) reviews some of this work, noting that such uses of machine translation are characteristic of experimental work in poetry and the digital humanities in general. She contrasts this playful approach with that of the machine translation developers who have recently turned their attention to the translation of literature, with the serious ambition of creating automatic systems whose output can compare with that of human translators. This latter branch of research, known as “literary-adapted machine translation” (Toral and Way 2015), is addressed below. We note here that, despite considerable interest in the ethical import of voice in human translation and, more recently, machine translation, there has been little work so far on the human translator’s voice in the machine translation of literary texts. This is a gap that we begin to address in this article. Human translators’ voices are present in the data used to train machine translation engines (see Kenny 2011a). They are also present in edits they make to machine-translated text, in comments they make in or on the text itself, and in their reflections on their role as research participants. In this article we focus on edits and, to a lesser extent, comments, as manifestations of voice. Before elaborating on our research design, however, we need to introduce a number of other distinctions made in the literature.

Textual and contextual voice

In Alvstad et al.’s (2017) voices framework, if a protagonist’s discursive presence is felt within a translated text, it is said to constitute a ‘textual voice’; if it is expressed using other means, for example in an introduction to a translation, a commentary on a translation, or a review of translation, it is said to constitute a ‘contextual voice’ (ibid.). A translator’s voice can thus be both textual and contextual, if, for example, as well as translating, they comment on their translation, or some aspect of the source text, in a footnote or elsewhere. Alvstad et al. (2017, 5-6) acknowledge common ground between what they call ‘contextual material’ and Genette’s (1987) ‘paratext’ (see also Taivalkoski-Shilov and Koponen 2017), but stress that contextual material has broader methodological scope. It can include, for example, comments made by a translator that are intended not to link between translated texts and their readers, but to address researchers studying a translation, as happens in our study. Any study of the translator’s textual voice needs a way of operationalizing the notion of voice. As indicated above, the translator’s textual voice can be induced from

1 Taivalkoski-Shilov (2013) similarly refers to ‘intratextual’ and ‘extratextual’ voices.
collections of the translator’s work in which consistent stylistic preferences can be discerned. Such stylistic preferences can, for example, be detected in the shifts that the translator typically makes vis-à-vis source texts (e.g. Pekkanen 2010). They may also stand out against the background of contrasting choices made by a second translator (or multiple other translators) of the same source text into the same target language (e.g. Winters 2005; Pekkanen 2010; Kolb 2017; Wang and Li 2020), or become salient because they are consistently observed in the translator’s work, no matter what source texts they are translating (Pekkanen 2010; Wang and Li 2020). In the context of corpus-based translation studies, Gabriela Saldanha has given a comprehensive definition of ‘translator style’ that serves us well in the current context as an operationalization of the translator’s textual voice. Translator style is, according to Saldanha (2011, 31):

   a ‘way of translating’ which:
   • *is felt to be* recognizable across a range of translations by the same translator,
   • distinguishes the translator’s work from that of others,
   • constitutes a coherent pattern of choices,
   • is ‘motivated’ in the sense that it has a discernible function or functions, and
   • cannot be explained purely with reference to the author or source-text style, or as the result of linguistic constraints.
   (Saldanha 2011, 31; our emphasis)

Similar to Folkart (1996), who defines voice as “a cluster of textual features that gives the impression of being attributable to a single source of enunciation” (Folkart 1996, 127; as translated by Taivalkoski-Shilov 2019, 695; our emphasis), Saldanha acknowledges a subjective element in the recognition of translator style. There has to be a subjectivity that ‘feels’ or that ‘forms an impression’ about a translator’s style. Different readers will form different impressions, and a textual feature that is a clear manifestation of a particular translator’s style to one reader might be overlooked by another. A translation scholar who has studied a particular translator in depth, perhaps

---

2 Folkart’s original reads “‘Grosso modo … une voix, c’est un faisceau de traits textuels qui donne l’impression de pouvoir être reconduit à un seul foyer d’énonciation’ (Folkart 1996, 127, note 3).
using vast quantities of the translator’s output to inform their analyses, but also interviewing the translator or conducting other kinds of ethnographic observation, might be expected to have particular insight into that translator’s style, insight that is, however, crucially underpinned by textual evidence. To that analyst, the translator’s style, and hence textual voice, might be more noticeable than it is to other readers of the translator’s work. In Greenall’s (2015) term, the translator’s voice is ‘manifest’ to the analyst. Analysts who have conducted such in-depth analyses of particular translators’ work include Winters (e.g. Winters 2015) and Borg (2016), whose studies we return to below.

Finally, we also draw here on Hermans’ (1996) distinction between the translator as a textual presence (i.e. a ‘voice’) and the real person, or “Biological Translator”. While originally put forward as a useful narratological clarification, the term biological translator takes on new significance in the context of machine translation workflows. The biological translator at the heart of our study is Hans-Christian Oeser, whose work is described later in this article.3

The translator’s voice in recent approaches to the machine translation of literature

For much of the history of machine translation, perhaps most of its literary uses were of the playful kind described above, where nonsense and noise were seen as aesthetic virtues, and machine translation output was used by creative writers. The transition from rule-based to data-driven machine translation that started in the late twentieth century and gathered pace in the 2000s however, saw the quality of machine translation output increase. At the same time, more and more data, including data from previously translated literary texts, became available to train what Toral and Way (2015) called ‘literary-adapted machine translation’ engines. The protagonists changed: computer scientists became more involved; and the aims changed: serious attempts were now made to train statistical machine translation (SMT) engines (e.g. Besacier 2014; Toral and Way 2015), and later neural machine translation (NMT) engines (Toral and Way 2018; Moorkens et al. 2018), that would produce accurate and fluent literary translations. Space restrictions prevent us from giving full details of

---

3 Many commentators now distinguish between ‘human translation’ and ‘machine translation’ in English. An analogous distinction is increasingly made in French between biotraduction and traduction automatique (see, e.g., https://tq2020.sciencesconf.org). While these terminological distinctions are necessary in some contexts, we take human translation as the default, which makes the qualifying adjective ‘human’ in the phrase ‘human translation’ superfluous in many other contexts.
these studies. Rather we wish to focus here on where and how translators’ voices have been present and/or manifest in this research.

In Besacier (2014), input from a literary translator came only at the very end of a long process involving the incremental training of an SMT engine, output of the machine-translated text, one round of post-editing, two rounds of revision, and an evaluation by nine readers. In this case, Jean-Yves Pellegrin, the source text’s usual translator into French, was also asked to evaluate the final product, and specifically whether the post-edited, revised text could serve as a starting point for a professional literary translator. His instinctive answer was ‘no’, albeit with some qualification.

In Toral and Way (2015), a literary translator’s voice is buried somewhere in the data used to train a literary-adapted SMT engine, but that translator is not named, and their voice does not become manifest. In subsequent research, Toral and Way (2018) compare the quality attained using NMT on literary text to that reached using phrase-based SMT on the one hand, and human translation on the other. This time a large corpus of 133 novels translated from English into Catalan is used to train both the SMT and NMT systems. Few qualitative details are given about the training data, and no details are given about the translators involved. There is no involvement from literary translators in using the engine or evaluating its output, although the authors express a desire to involve professional literary translators in future research.

Toral et al. (2018) use the same two machine translation engines as described in Toral and Way (2018) to translate a chapter from a fantasy novel into Catalan. This time they ask six professional translators with previous experience in literary translation to translate subsets of this chapter under three alternating conditions: ‘from scratch’, post-editing of SMT, and post-editing of NMT. They record the time taken to translate each sentence, all keystrokes, and the number and duration of pauses for each translator, in an attempt to measure the temporal, technical and cognitive effort involved in translating/post-editing in the three conditions. On this occasion, four out of the six literary translators are named in the acknowledgements in their capacity as research participants (two choose to remain anonymous), although participants are, quite properly, anonymized in the body of the study.

Moorkens et al. (2018), in turn, use data gathered during the same study to examine the six participant translators’ perceptions of the task. The study then contrasts these perceptions with the measurements of temporal and technical effort captured in Toral et al. (2018). Moorkens et al. (2018) give information about their translator-
participants’ qualifications and their, fairly modest, experience as post-editors, as well as their experience in literary translation. They capture data on participants’ perceptions of the task in a post-task questionnaire and debriefing conversations. They are thus in a position to reproduce literary translators’ opinions about machine translation in general and the task just completed in particular, and hence to give limited voice to those translators, if voice is understood as contextual voice of the kind “generated during the research process” (Alvstad et al. 2017, 5). Although Moorkens et al. (2018) is probably the high point so far in the history of investigations of literary-adapted machine translation from the point of view of the visibility and audibility of the literary translator, what is still missing is any consideration of the translator’s textual voice in the context of machine translation workflows.

The experimental settings used in the studies reviewed here also need to be problematized. They usually involve participants coming to the researchers’ lab and using a dedicated post-editing interface that segments source and machine-translated text into aligned, often sentence-based units, much like a translation memory tool would. Although there are exceptions (see, e.g. Taivalkoski-Shilov 2019, 691; Rothwell 2020) most literary translators work from home however, and use standard word processing applications like Microsoft Word, so the experimental settings used to date in literary-adapted machine translation studies cannot really show us what it might be like for a literary translator to post-edit a machine-translated text in their normal working environment.

Towards more ecologically valid approaches to the study of literary translators at work

Models for more ecologically valid studies come not from literary-adapted machine translation, but from translation-oriented workplace research (Risku et al. 2017). One of the few researchers to carry out such workplace research with literary translators is Waltraud Kolb, who conducted an experiment in which five (unnamed) translators were asked to translate a short story into German. Kolb (2017) designed her experiment so that the translators worked in their normal workplace, usually at a desk in their home. The translators used Translog to capture their keyboard activity and Audacity to audio-record concurrent and retrospective verbalizations as and after they completed their translations.
In a related study, Claudine Borg (2016) used direct observation of a single literary translator at work, in this case the highly accomplished translator into Maltese, Toni Aquilina. As well as her own field notes, she gathered data from Aquilina’s verbalizations, using audio or video-recordings to fill gaps in her field notes, and from retrospective and semi-structured interviews. She also had at her disposal all drafts of the translation completed by Aquilina. In deference to her research participant’s preferences, she did not use a keyboard-logging tool (2016, 70).

Kolb and Borg thus collect data that can shine a light on the translator’s textual and contextual voice, a central concern in both studies. Voice, is however, just one of their concerns, as both researchers seek to explore a wide range of phenomena associated with the translation process. They thus select a suitably wide range of data collection techniques available to the ethnographically-inclined translation scholar, while accommodating their participants’ preferences. We, on the other hand, are less broadly concerned with the translation process in our study. Rather we are motivated by a specific desire to trace the translator’s textual and contextual voice in a workflow that involves the post-editing of a machine-translated literary text. Like Kolb and Borg, we are committed to conducting our research in as ecologically valid a way as possible, and in a way that allows our research participant’s voice to be heard, but given our focus on textual voice as a kind of post-hoc trace of the translator’s presence, we tend more towards product-oriented than process-oriented analysis. We thus eschew the kind of observational techniques used by Kolb (2017) and Borg (2016), for example video-recording or keyboard logging. This, arguably, increases the ecological validity of our research, as an authentic work setting for a literary translator would not normally involve the use of these techniques, despite their importance in experimental research. On designing the study outlined below, our expectation was that we would find evidence of the translator’s textual voice in post-edits he would make (or not) to a machine-translated text. We expected to hear his contextual voice in comments he might write as he post-edited. In the following section we outline our research design in detail.

Research Design

Participant
As indicated above, the literary translator with whom we worked is Hans-Christian Oeser. Oeser is a translator and editor of international significance, and probably the most important translator of Irish literature into German (Winters 2015). In our view, he easily exceeds the threshold set for recognition as a professional or expert in sources such as Kolb (2017). Oeser’s translation work has been the focus of a number of studies by Winters (2004, 2007, 2009, 2010), who used corpus techniques to discover elements of his textual voice. Winters’ earlier treatment of Oeser’s style was informed by a comparative corpus-driven analysis in which his style could be offset against that of Renate Orth-Guttman in their respective translations into German, published in the same year (1998), of F. Scott Fitzgerald’s (1922) *The Beautiful and Damned*. Winters (2015) is based both on earlier corpus-driven work and an extensive 2014 interview with Oeser. On the basis of these sources, and drawing on Saldanha’s operationalization of ‘style’ reproduced earlier in this article, we can say that prominent features of Oeser’s style relevant to the current study are that he:

1. stays as close to the source text as possible. Oeser (in Winters 2015, 214) calls himself a *Wortwörtlichkeitsfanatiker*, or a fan of literalism, where such literalism is a means of revealing elements of the source text author’s style but does not make a virtue of disrupting the target language reader’s experience, unless of course Oeser is emulating a deliberately disruptive feature of the source text (ibid.)
2. favours subject-verb inversion
3. avoids the use of the optional complementizer *dass* (‘that’) in sentence complements, for example those introducing reported speech or thought
4. likes to use alliteration, and other poetic or lyrical elements

4 Over a career that has spanned four decades he has translated over 160 works of literature (novels, short stories and poetry). His numerous accolades include the Aristeion European Literary Translation Prize for *Der Schlächterbursche* (The Butcher Boy by Patrick McCabe) in 1997, the Heinrich Maria Ledig-Rowohlt Prize for his contribution to literary translation in 2010, and multiple mentions in the most influential German-language ‘best books’ lists (e.g., SWF-Bestenliste, ORF-Bestenliste, KrimiZEIT-Bestenliste). Further information on Oeser’s career to date is available from Winters (2015) and Oeser’s personal website http://www.hanschristianoeser.com/.

5 The second position in main clauses in German is reserved for the finite verb. If the grammatical subject occupies the first position the canonical subject-verb-complement order applies. If, however, something other than the subject takes first position, for example an adjunct, subject-verb inversion occurs, as the finite verb still needs to occupy the second position.

6 The use of such a complementizer in German might, however, reflect an element of the style of the source language author, in which case Oeser might prioritize retention of the author’s style. For an example, see Kenny (2006, 52-53; reproduced from Winters 2005).
is committed to preserving lexical richness in translation, and argues that “translations cannot be good if you only have one word at your disposal where you should have ten” (Winters 2015, 213)

- Pays close attention to register and very consciously uses higher register when appropriate

Oeser is a literary translator with whom one of the authors has thus engaged as a researcher over many years. This mitigates some of the practical challenges in ethnographic research (to which our research is roughly aligned) considered by Risku (2017), as we were already in touch with an “active and motivated contact person in the empirical setting” (ibid. 296). We sought, and gained, ethics approval from Heriot-Watt University before conducting our empirical research. From the outset, there was no question of keeping Oeser’s identity confidential, given the nature of the study we planned to conduct. This made the issue of informed consent all the more important.

**Experiment set up**

The machine translation system selected for use in our experiment was DeepL, as Oeser judged it to have reached a quality threshold that made it potentially interesting for many kinds of translators, even for what had formerly been considered particularly difficult language pairs like English-German. DeepL had the merit of high credibility, accessibility and usability. Although we could not customize the system for our purposes, customizability was less of a concern for us on this occasion as we were aiming for ecological validity and deemed DeepL to be the system that would most likely be chosen by Oeser if he had chosen to initiate the trial himself. While Oeser was not enthused by ‘fictitious briefs’ (Borg 2016, 60), he agreed to accept such a brief from us for the sake of the research. The brief would involve post-editing a

---

7 To this extent, we had similar concerns to Borg (2016), and many of the ethical issues she considers (ibid. 67-68 and Appendix 1) also apply to our work.
machine-translated version of (part of) a text he had previously translated from scratch, but no further details were given.  

We subsequently chose an excerpt from F. Scott Fitzgerald’s *The Beautiful and Damned*, translated it using DeepL Pro, and sent the machine translation, along with the source text, in Microsoft Word format to Oeser by email for post-editing. Word format was chosen as this was the editing environment that Oeser would use naturally for a translation/editing/revision job, and all his edits and comments could be captured conveniently using the track changes and commenting functions. We also sent him the full source-text in electronic form, so that the excerpt could be contextualized. We explained the aim of the experiment, which was to investigate how a literary translator uses a commercial machine translation program (DeepL Pro) in his normal working environment under normal working conditions. We explained that we would compare the translation resulting from this experiment with his translation from 1998, which was produced without any machine translation aid. As agreed from the outset – and because neither productivity nor the translation process *per se* were the focus of our research – we did not ask Oeser to record how long it took him to post-edit the excerpt, and we did not observe him at work. The brief given to him was to produce a translation of the source text based on the machine-translated output, one that would be ready for submission to the publisher, and would meet Oeser’s usual quality criteria. We also asked Oeser not to refer to his previous translation, but to note down any thoughts that occurred to him during the translation process, and that he wanted to share with us, as long as this did not constrain him unnecessarily or compromise his normal way of working. Given that the task was not being timed, this would not interfere with our results.

The excerpt we chose for the experiment comprises the first chapter of Book II of *The Beautiful and Damned*, entitled ‘The Radiant Hour’. It is 1,370 words long and contains 92 sentences. It was selected because Oeser’s original from-scratch translation of this chapter, as investigated in Winters (2005), had a particularly high density of features considered typical of his style. We expected Oeser to have some recollection of translating the book more than twenty years previously, but not to the extent that he would be able to deliberately retrieve from memory exactly what he had

---

8 It is clear from Oeser’s later reaction, that if he did anticipate receiving a particular text, he anticipated incorrectly.
written before. In any case, it was clear to Oeser that the aim was not to recreate his previous translation. Rather the brief was to bring the machine-translated text up to his normal standard, to make it into ‘an Oeser’.10

Research Questions
The general questions that we intended to answer in this study were as follows:

- Is Oeser’s textual voice discernible in his post-edited translation?
- If so, to what extent?
- How does Oeser use his contextual voice in this experiment?

Research Materials
The research materials we had at our disposal to answer these questions were:

1. The source text (labelled ‘FSF’ below)
2. Oeser’s from-scratch translation of the excerpt in question (‘HCO 1998’)
3. The machine translation output (‘MT’)
4. Oeser’s post-edited version of the machine translation output (‘HCO 2019’)
5. Oeser’s comments on the task as recorded using the comments function in Microsoft Word
6. The transcript of a follow-up interview11
7. Previous studies of Oeser’s style as listed above.

Procedure
Texts 1 to 4 above were first aligned at sentence level using a table structure in Microsoft Word. The principal analysis of textual voice was based on manual inspection of this table.
In the first instance we were interested in:

- What Oeser retained from the machine translation.
- What Oeser changed in the machine translation.
- Where he retained elements, whether such retentions made the post-edited text more or less consistent with his 1998 translation/known elements of his style.

---

9 It was established in the post-task interview that Oeser did not recall the details of his original translation.
10 In German: das zu ‘einem Oeser’ zu machen.
11 Space constraints prevent us from elaborating on the post-task interview in the current article.
Where he made changes, whether such changes made the post-edited text more or less consistent with his 1998 translation/known elements of his style.

Our identification of units of analysis drew on Toury’s (1995) ‘coupled pairs’ approach, which involves the identification of problem-solution pairs on the basis of their relevance to the analysis at hand.\(^\text{12}\) An example from the text should help elucidate the process. The first sentence in the excerpt reads as in Example 1a:

\[\text{Example 1a (FSF)}\]
\[
\text{After a fortnight Anthony and Gloria began to indulge in ‘practical discussions’, as they called those sessions when under the guise of severe realism they walked in an eternal moonlight.}
\]

The machine-translated version appears in Example 1b:

\[\text{Example 1b (MT)}\]
\[
\text{Nach zwei Wochen begannen Anthony und Gloria, sich in "praktischen Diskussionen" zu vergnügen, wie sie diese Sitzungen nannten, als sie unter dem Deckmantel des strengen Realismus in einem ewigen Mondlicht wandelten.}
\]

Oeser’s post-edited version is as follows:

\[\text{Example 1c (HCO2019)}\]
\[
\text{Nach zwei Wochen begannen Anthony und Gloria, sich in „praktischen Diskussionen“ zu ergehen. So nannten sie jene Sitzungen, da sie unter dem Deckmantel eines strengen Realismus in ewigem Mondlicht wandelten.}
\]

The first edit that Oeser makes here is to the punctuation, replacing inappropriate English inverted commas with German ones. He then changes the incorrect lexical item \textit{vergnügen} to \textit{ergehen}.\(^\text{13}\) He goes on to split the long machine-translated sentence into two, and links his sentences with \textit{So}, a conjunctive adjunct roughly equivalent to ‘this is how’, and that replaces \textit{wie} (‘as’). This, in turn, requires a change of word order, as the finite verb \textit{nannten} (‘called’) moves forward to occupy the second position in what is now a main clause. Oeser changes \textit{diese} (‘these’) to \textit{jene} (‘those’) following the source-text usage of the distal form. He then changes \textit{als} to \textit{da}, as \textit{als} is

---

\(^{12}\) For an evaluative summary of Toury’s approach, see Kenny (2011b: 80-82).

\(^{13}\) \textit{vergnügen} means to ‘enjoy’, but does not impart the slightly disapproving tone of ‘indulge’. It also does not call for the strongly-bound preposition \textit{in} as used in the machine translation; and it does not quite collocate with \textit{Diskussionen}. 
not an appropriate translation of ‘when’ in the case of habitual iteration in the past. Here \textit{wenn} would be appropriate, but Oeser prefers the higher-register \textit{da}. He changes \textit{des} (‘of the’) to \textit{eines} (‘of a’) and deletes \textit{einem} (‘a’ in the dative case) so that \textit{Mondlicht} (‘moonlight’) becomes, in line with convention, an uncountable noun in German. Oeser thus makes at least nine changes to the first machine-translated sentence, but most of these changes are required either to produce grammatically correct German, or to fix undeniably incorrect translations. We can characterize such edits as arising from ‘linguistic constraints’ (Saldanha 2011 and above). While they are thus of interest to researchers who study post-editing from other – especially productivity – perspectives, they are of less interest in our study, which is concerned with translator style.\footnote{Our approach is in stark contrast to the approach usually adopted in post-editing research, where the analyst is interested in the errors post-editors (are forced to) fix as well as the changes they make simply to bring the text into line with what are their own preferences. We note however, that although we do not consider Oeser’s fixing of errors as a manifestation of his textual voice, we do consider his comments on some of these errors as a manifestation of his contextual voice.} Of all the edits that Oeser makes in Example 1c, it is perhaps the use of the temporal conjunction \textit{da} that is most revealing of his style as evidenced in previous studies, as it is a self-conscious use of a higher-register option offered by the target language. Coincidentally, it is the only translation decision Oeser makes in the sentence in question that he also comments on:

\begin{center}
\begin{quote}
\textit{„als“ ist falsch, „wenn“ richtig, weil iterativ. Bei mir „höherer Stil“: „da“.}
\end{quote}
\end{center}

(‘als is wrong; wenn correct, because iterative. I’ll go for higher register: da.’)

Here textual voice chimes with contextual voice, and the edit is recorded as one that speaks to the presence of Oeser’s style and thus his textual voice in the post-edited text.

In a second phase of analysis, the post-edited sentence (1c), repeated here for convenience, is compared with Oeser’s (1998) from-scratch translation (1d):

\begin{center}
\begin{quote}
Example 1c (HCO2019)
Nach zwei Wochen begannen Anthony und Gloria, sich in „praktischen Diskussionen“ zu ergehen. So nannten sie jene Sitzungen, da sie unter dem Deckmantel eines strengen Realismus in ewigem Mondlicht wandelten.
\end{quote}
\end{center}

\begin{center}
\begin{quote}
Example 1d (HCO1998)
\end{quote}
\end{center}
Nach vierzehn Tagen begannen Anthony und Gloria sich in »praktischen Diskussionen« zu ergehen; so nannten sie es, wenn sie zusammensaßen und hinter der Maske strenger Vernunft in ewigem Mondenschein wandelten.

Here we note that *Nach zwei Wochen* (‘after two weeks’), which was retained from the machine translation output, is at odds with Oeser’s from-scratch translation *Nach vierzehn Tagen* (‘after fourteen days’), but we do not consider this difference as particularly indicative of either consistency with or departure from Oeser’s style; rather it is an example of two synonymous expressions that can alternate without affecting style or meaning. The change that Oeser made to the machine translation output whereby he split one long sentence into two, the second starting with *So*, is consistent with the approach he took in his from-scratch translation twenty years previously, but it does not constitute a feature of Oeser’s style in our current analysis.

A marked difference between (1c) and (1d) is the use of the nominal *Sitzungen* (‘sessions’) in the former, and the verbal *wenn sie zusammensaßen* (‘when they sat together’) in the latter. The complementary distribution of German nominal and verbal structures is an interesting topic (see Bisiada 2018) but, again, we have not associated a particularly nominal or verbal style with Oeser in the past, and do not, for current purposes, consider this difference as indicative of a departure from his style. It is interesting to note, however, that the temporal conjunction used by Oeser in his 1998 from-scratch translation is *wenn*, which is the form he considered using in his 2019 post-edited version, but ultimately rejected in favour of higher-register *da*. This is an example of Oeser editing the machine translation output in a way that is not literally consistent with his 1998 translation (that is, he has not chosen the exact same word), but that is nonetheless functionally consistent with his style as identified in previous studies (as he has chosen a higher-register option).

There follow three noun-pairs, the members of which differ between Examples 1c and 1d. *Deckmantel* (‘cover’) and *Maske* (‘mask’), and *Realismus* (‘realism’) and *Vernunft* (‘reason’), can be analysed as synonymous pairs (in this context anyway) that do not necessarily speak to Oesers’s style. The pair *Mondlicht* (‘moonlight’) and *Mondenschein* (‘moonshine’) is, however, interesting from our point of view, as the latter term is clearly poetic and reflects Oeser’s predilection for such forms. In this case, by retaining the more neutral *Mondlicht* is his post-edited version, Oeser seems to have ‘dialled down’ his style in terms of register; but this observation needs to be juxtaposed with the finding that Oeser simultaneously ‘dialled up’ his style in the
same version by using the high-register *da* earlier in the sentence. The above example, it is hoped, demonstrates the complexity of the analysis. It relies not just on comparing words and structures in three German versions of the chapter in question, but on evaluating the potential of those words and structures for expressing Oeser’s style. At the risk of presenting a reductionist view of our own analysis, we can say, in summary, that in the first sentence of the excerpt in question, Oeser introduces one style-intensifying edit to the machine translation output (in *da*) but misses the opportunity to introduce another possible style-intensifying edit (by not using *Mondenschein*). The first observation comes from the systematic comparison of his post-edits (in HCO2019) with the machine translation output (MT); the second observation from the systematic comparison of his post-edits (again HCO2019) with his from-scratch translation (HCO1998). These observations, which seem to cancel each other out, nevertheless begin to answer the general research questions presented above: Oeser’s textual voice *is* discernible in the post-edited sentence in question, and it is discernible to the same extent as in his from-scratch translation, although it is not necessarily instantiated in the same words in the two texts. In other words, traces of his style are distributed across different parts of the two sentences. In both cases, his comments inform our analysis, as does previous knowledge of his style. As to the question of how Oeser uses his contextual voice, we note that, in this example, he uses it to make a self-reflexive comment about his style – in this case his preference for higher-register forms.

In the following section, we present summary quantitative results from an initial analysis of textual voice in the entire excerpt. We consider what we deem to be style-relevant retentions of, and style-relevant changes to, the machine translation output. We counted as retentions those lexical items and syntactic structures/instances of word order that were the same in the machine translation (MT) and in Oeser’s post-edited version (HCO2019), but which differed from his from-scratch translation (HCO1998). Style-relevant retentions were thus the subset of retentions that could be said to either amplify or diminish Oeser’s style, as exemplified by *Mondlicht* in the discussion under Example 1 above. We counted as changes those lexical items and syntactic structures/instances of word order that were different in the machine translation (MT) and in Oeser’s post-edited version (HCO2019). Style-relevant changes are thus the subset of changes that could be said to either amplify or diminish Oeser’s style, as exemplified by *da* in the discussion under Example 1 above.
In order to give more qualitative insight into Oeser’s post-edits, we drill down into the data to present findings for three features that are typical of Oeser’s style, namely a preference for inversion, use of colloquial language in dialogue and use of higher-register lexis. All coding of units as belonging to these categories was done on the basis of consensus between two native-German speaking evaluators. Finally, the analysis of contextual voice is based on the comments that Oeser made in his post-edited text.

**Analysis**

We identified 74 style-relevant instances in Oeser’s post-edited translation that relate to the use of (i) inversion, (ii) colloquial language and (iii) register. In 36 cases, these style-relevant instances were retentions from the machine translation that caused some kind of diminution of Oeser’s style as evidenced in his 1998 translation (Table 1). In 38 cases, these style-relevant instances were changes to the machine translation that amplified Oeser’s style (Table 2).

<table>
<thead>
<tr>
<th>Feature</th>
<th>No. of instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of inversion</td>
<td>14</td>
</tr>
<tr>
<td>Less colloquial language</td>
<td>11</td>
</tr>
<tr>
<td>Lower register</td>
<td>11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>36</strong></td>
</tr>
</tbody>
</table>

Table 1. Style-diminishing retentions: analysis of selected features

Table 1 indicates that, in 36 instances, by following the machine translation, Oeser has (i) missed the opportunity to use inversion, (ii) used less colloquial fictional dialogue, or (iii) used a lower register in his post-edited text. Example 2c below illustrates lack of inversion. Example 3c illustrates the use of less colloquial language than in 1998. The use of lower register (in *Mondlicht*) has already been illustrated in Example 1c above.

Example 2c is a case where Oeser repairs an error in the machine translation and makes lexical changes, but maintains the subject-verb order of the machine translation
output, despite having approached the sentence differently in 1998, where he used an inversion. 15

Example 2a (FSF)
He was a relief from younger men…

Example 2b (MT)
Er war eine Erleichterung für jüngere Männer…
(‘He was a relief for younger men…’)

Example 2c (HCO2019)
Er war eine willkommene Abwechslung von den jüngeren Männern…
(‘He was a welcome change from the younger men…’)

Example 2d (HCO1998)
Nach all den jüngeren Männern war er eine Wohltat…
(‘After all the younger men he was <inverted> a relief…’)

Example 3c shows retentions of both a syntactic and lexical nature. Here, a subordinate clause introduced by the complementizer dass is retained by Oeser, and the less colloquial glaube is retained in preference to finde, as used in Oeser’s from-scratch translation.

Example 3a (FSF)
I don’t think that’s right.

Example 3b (MT)
Ich glaube nicht, dass das richtig ist.
(‘I don’t believe, that that is right.’)

Example 3c (HCO2019)
Ich glaube nicht, dass das richtig ist.
(‘I don’t believe, that that is right.’)

Example 3d (HCO1998)
Ich finde das nicht richtig.
(‘I find that not right.’)

The influence of the machine translation as described in Table 1 and examples above is balanced out, however, by almost the same number (38) of changes to the machine translation that are consistent with Oeser’s style (Table 2). Nevertheless, only 17 of

15 German examples are presented with a literal but readable ‘back translation’ designed to elucidate selected linguistic features of the example. Angle brackets are used in cases where further metalinguistic comment is required.
these changes are ‘new’ manifestations of Oeser’s style in the post-edited translation (Column 3). The remaining 21 changes (Column 4) are ‘renewed’ manifestations of his style as these edits result in the same translation as in 1998.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Inversion</td>
<td>13</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>More colloquial language</td>
<td>12</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>Higher register</td>
<td>13</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>38</strong></td>
<td><strong>17</strong></td>
<td><strong>21</strong></td>
</tr>
</tbody>
</table>

Table 2. Style-amplifying changes: detailed analysis of selected features

Example 4c shows Oeser introducing an inversion and more colloquial language into his post-edited text, both of which constitute new manifestations of his style.

Example 4a (FSF)
I think she’s cute…

Example 4b (MT)
Ich finde sie süß…
(‘I find her sweet…’)

Example 4c (HCO2019)
Die finde ich süß…
(‘Her <informal> I find<inverted> sweet…’)

Example 4d (HCO1998)
Ich finde sie niedlich…
(‘I find her cute…’)

In Example 5c, Oeser uses higher register in his post-edited text than he did in his translation from scratch, and this is again understood as a new manifestation of his style:

Example 5a (FSF)
…within a month he had asked her to marry him…

Example 5b (MT)
…innerhalb eines Monats hatte er sie gebeten, ihn zu heiraten…
(‘…within a month he had asked her to marry him…’)

Example 5c (HCO2019)
…innerhalb eines Monats hatte er um ihre Hand angehalten…
(‘…within a month he had asked for her hand…’)

Example 5d (HCO1998)
…binnen eines Monats machte er ihr einen Antrag…
(‘…within a month he made her an offer…’)

Examples 6c, 7c, 8c and 9c, on the other hand, contain renewed manifestations of Oeser’s style, as they illustrate the same translation-stylistic behaviour as was manifest in his 1998 from-scratch translation.

In Example 6, Oeser uses the same inversion in his post-edited translation (6c) and in his 1998 from-scratch translation (6d). In this case, the inversion accompanies the fronting of ‘Rachael’, which ensures that contrastive emphasis is put on this name. Oeser comments on the emphatic function of fronting in this case (Comment 2), but we also note that the move results in a stress pattern that better approximates spoken German and is thus simultaneously a change that increases the colloquial nature of the dialogue in his post-edited text.

Example 6a (FSF)
...but I rather like Rachael.

Example 6b (MT)
...aber ich mag Rachael lieber.
(‘...but I like Rachael better.’)

Example 6c (HCO2019)
...aber Rachael mag ich ziemlich gern.
(‘...but Rachael I like quite a lot.’)

Example 6d (HCO1998)
...aber Rachael mag ich ziemlich gern.
(‘...but Rachael I like quite a lot.’)

Comment 2
Betonung, also nach vorne.
(‘Emphasis, so move to the front.’)

In Example 7c, he drops the dummy subject Es, making the dialogue more colloquial, just as it was in 1998 (7d).
In Example 8c Oeser uses the higher-register *Friede* (‘peace’), exactly as he had done in 1998 (8d).

Example 8a (FSF)
Peace was restored…

Example 8b (MT)
Der Frieden wurde wiederhergestellt…
(‘The peace was<eventive> restored…’)

Example 8c (HCO2019)
Der Friede war wiederhergestellt…
(‘The peace was<stative> restored…’)

Example 8d (HCO1998)
Der Friede war wiederhergestellt…
(‘The peace was<stative> restored…’)

He also comments on this form in his post-edited text, remarking that:

Comment 3
Friede ohne n klingt altertümlicher und schöner.
(‘Friede without the ‘n’ sounds more old-fashioned and nicer.’)

In another case, Example 9, Oeser replaces *warum* (‘why’) with the less frequent, and in our opinion, higher-register form *weshalb* (‘why>):

Example 9a (FSF)
Then tell me all the reasons why you're going to marry me in June.
Example 9b (MT)
Dann sag mir alle Gründe, warum du mich im Juni heiraten wirst.
(‘Then tell me all reasons why you will marry me in June.’)

Example 9c (HCO2019)
Dann nenne mir all die Gründe, weshalb du mich im Juni heiraten wirst.
(‘Then name all the reasons why you will marry me in June.’)

Example 9d (HCO1998)
Dann zähle mir sämtliche Gründe auf, weshalb du mich im Juni heiraten willst.
(‘Then enumerate all the reasons why you want to marry me in June.’)

In this case, Oeser makes another self-reflexive comment, explaining the edit solely in terms of his style (Comment 4):

Comment 4
Auch so eine Marotte von mir: lieber "weshalb" als "warum" zu verwenden.
(‘Another one of my quirks: I prefer using weshalb to warum.’)

To sum up, in the comparison of Oeser’s post-edits with the machine translation output, we note an almost equal number of style-diminishing retentions (36, Table 1) and style-amplifying changes (38, Table 2). While the machine translation output appears to be priming Oeser to accept some solutions that are inconsistent with his style, he appears to correct this tendency in almost equal measure through the changes he makes to other solutions. However, these changes result in only 17 (Table 2, Column 3) additional manifestations of his style in his 2019 post-edited translation, while 21 (Table 2, Column 4) of these changes result in renewed manifestations of his style, which illustrate the same translation-stylistic behaviour as in his 1998 from-scratch translation. Therefore, the comparison of his post-edited translation with his 1998 from-scratch translation reveals that despite those changes, his post-edited work is still somewhat less representative of his style than his 1998 from-scratch translation. If the features we singled out for closer analysis can be taken as indicators of Oeser’s style, then a crude calculation based on the data presented in Tables 1 and 2 above suggests that instances of Oeser’s style are less present by approximately a third in his post-edited work than in his translated work. His style, and thus his textual voice, is still discernible in the edits he makes to the machine translation output, but the overall effect is perhaps a slight dampening of his textual voice.
Oeser’s contextual voice is present in the 73 comments that he made alongside the hundreds of edits to the machine translation output. The majority of these are of a metalinguistic nature, highlighting particular challenges posed by the source text, or problems generally faced by the translator into German, or pointing out weaknesses in the machine translation. These demonstrate Oeser’s stylistic awareness. Numerous comments relate to choices that affect the aesthetic qualities of the emerging target text, such as register, alliteration and lexical preferences. Although space constraints prevent us from discussing his comments at length, we note here that Oeser’s comments frequently reinforce his textual voice and thus our decisions about what constitutes his style, as illustrated in the comments reproduced above. His comments also represent a treasure trove of observations on the particular difficulties – for example, those related to tense and cohesion – that creative narratives present for machine translation, and could contribute to the nascent body of research that looks at these and other similarly demanding text types (e.g. Poncharal 2020).

**Conclusion**

In this article we hope to have reaffirmed the ethical import of voice in literary translation, and drawn attention to the scant attention the translator’s voice has received thus far in discussions of literary-adapted machine translation. We have problematized the experimental settings used in recent studies of literary-adapted machine translation, and made the case for a more ecologically valid approach. Against this background we have presented an initial study that attempts to investigate whether one translator’s textual voice is manifest in his post-editing work and how it has been influenced by the use of machine translation. We also aimed to capture the translator’s contextual voice in a way that has not previously been discussed, to our knowledge, in the literature. The line-by-line analysis of textual voice and comment-by-comment analysis of contextual voice that we propose is labour-intensive yet sensitive enough to detect contrasting style-amplifying and style-diminishing edits and non-edits. On the basis of our initial analysis of textual voice, we tentatively conclude that the translator’s voice is somewhat dampened in his post-editing work. We note, however, how he uses his contextual voice to justify his translation/post-editing choices by referring to source-language challenges, machine translation errors, narrative structures in German, and – crucially – elements of his
own style. His contextual voice thus plays a supporting role in our analysis of his
textual voice. It also provides insight into the various challenges that machine
translation systems still face in dealing with literary texts.
Finally, the initial analysis presented here provides just a partial glimpse into the rich
set of data elicited in the experiment on which we report. In future work we hope to
integrate more features of Oeser’s textual voice, and to take a deeper look at his
contextual voice, as attested both in his comments on his post-edited text and in the
post-task interview. We also intend to work with larger samples and non-fictitious
briefs. The present article has, we hope, laid the groundwork for such an extended
study.

Funding Information
The research reported on in this article benefited from an Internal Research Grant
from the School of Social Sciences at Heriot-Watt University, Scotland, and a travel
grant awarded by the Faculty of Humanities and Social Sciences, Dublin City
University, Ireland, under its Spring 2019 Faculty Conference Travel Scheme.

Acknowledgements
We are extremely grateful to Hans-Christian Oeser for agreeing to participate in the
research reported on in this article. We would also like to thank the anonymous
reviewers who provided us with invaluable feedback.

References
Alvstad, Cecilia, Annjo K. Greenall, Hanne Jansen, and Kristiina Taivalkoski-Shilov
(eds). 2017. Textual and Contextual Voices of Translation. Amsterdam/Philadelphia:
John Benjamins.
Translator.” Target 12(2): 241-266.
Edited by Michael Holquist. Translated by Caryl Emerson and Michael Holquist.
Austin: University of Texas Press.
the 21ème Traitement Automatique du Langage Naturel (TALN 2014), Marseille.


Authors’ Addresses
Dorothy Kenny
http://orcid.org/0000-0002-4793-9256
Dublin City University
School of Applied Language and Intercultural Studies
Glasnevin
Dublin 9
Ireland
dorothy.kenny@dcu.ie

Marion Winters
https://orcid.org/0000-0003-4049-3156
Heriot-Watt University
Centre for Translation and Interpreting Studies in Scotland
Henry Prais Building 2.08
Edinburgh EH14 4AS
Scotland
m.winters@hw.ac.uk