

First Steps Towards the Creation of a Poetry Translation Mapping System

Burkhard Meyer-Sickendiek and Hussein Hussein and Timo Baumann

Abstract The translation of poetry is a complex, multifaceted challenge: the translated text should communicate the same meaning, similar metaphoric expressions, and also match the style and prosody of the original poem. Research on machine poetry translation is existing since 2010, but for four reasons it is still rather insufficient: A) The few approaches existing completely lack any knowledge about current developments in both lyric theory and translation theory; B) they are based on very small datasets; C) they mostly ignored the neural learning approach that superseded the long-standing dominance of phrase-based approaches within machine translation; and D) they have no concept concerning the pragmatic function of their research and the resulting tools. Our paper describes how to improve the existing research and technology for poetry translations in exactly these four points. With regards to a) we will describe the ‘Poetics of Translation’; with regards to b) we will introduce the Worlds largest corpus for poetry translations from *lyrikline*; with regards to c)

Burkhard Meyer-Sickendiek

Department of Literary Studies, Freie Universität Berlin, Berlin, Germany,

✉ bumesi@zedat.fu-berlin.de

Hussein Hussein

Department of Literary Studies, Freie Universität Berlin, Berlin, Germany

✉ hussein@zedat.fu-berlin.de

Timo Baumann

Language Technologies Institute, Carnegie Mellon University, Pittsburgh, USA

✉ tbaumann@cs.cmu.edu

ARCHIVES OF DATA SCIENCE, SERIES A
(ONLINE FIRST)

KIT SCIENTIFIC PUBLISHING

Vol. -, No. -, -

DOI 10.5445/KSP/XXXXXXXX/XX

ISSN 2363-9881



we will describe first steps towards a neural machine translation of poetry; with regards to d) we will describe first steps towards the development of a poetry translation mapping system.

1 Introduction

Although we know of a growing number of automatic poetry generation studies Oliveira (2012); He et al (2012); Yan et al (2013); Zhang and Lapata (2014); Yi et al (2016); Wang et al (2016); Hopkins and Kiela (2017); Oliveira (2017), to date there are just three publications existing in the field of automatic poetry translation. Greene et al (2010) used a phrase-based machine translation approach to translate Italian poems to English translation lattices, searching these lattices for the best translation that obeys a given metrical pattern. Genzel et al (2010) also use phrase-based machine translation to translate French poems to English ones, applying the meter and rhyme constraints during the decoding process rather than post-hoc analysis. Both methods report total failure in generating any translations with a fixed meter and rhyme format for most of the poems. Ghazvininejad et al (2018) presented a neural poetry translation system translating a French source poem to an English poem using a previous poetry generating approach Ghazvininejad et al (2017). None of these few approaches in the constantly growing field of machine translation fits to the challenges to produce viable translations of modern and postmodern poetry, basically because all these attempts are focused on meter and rhyme used mainly in traditional stanza types. By contrast, our paper focuses on translating free verse poetry, so instead of identifying the metrical feet (iamb, trochee) and the rhyme-pairs, our paper will offer first steps towards the automatic translation of poetry with regards to the rhythmical features being typical of modern and postmodern poetry.

The theoretical background of our paper is Henri Meschonnic's translation theory Meschonnic (1999). It offers one of the rare solutions to the greatest problem of poetry translation: the discrepancy between form and content. Criticizing the traditional translation of foreign languages in terms of a "langue", i.e. the individual language, Meschonnic called for translating the "discours", i.e. the concrete speech. The traditional attempts based on the sense of "langue" just translated the linguistic content from one language into another, potentially missing what has to be translated: the poetic text itself, not the foreign lan-

guage. In contrast, Meschonnic's translation focuses not only on what the text or its words say, but above all what they do, their orality, their rhythm, their performance, their voice. The voice as orality, as "the subject we hear" is for Meschonnic not a phenomenology or psychology of reading but something emerging from the text as "what a subject does to its language" (Meschonnic, 2011, pp. 139). This theoretical approach is based on linguist Emil Benveniste's notion of rhythm and discourse. Benveniste demonstrated that the original notion of rhythm developed by pre-Socratic philosophy (Democritus and Heraclitus) was understood not in the sense of a stable essence, but as the shape of a given instant, and that Plato has altered it profoundly by confounding the notion of rhythm with that of meter, establishing an objective, regular wave-like coming and going associated with metrics, that is imposed upon the individual form. Meschonnic's 'Poetics of Translation' is based on this pre-Platonic notion of rhythm reconstructed by Benveniste. This impact lead to a "performative turn" in translation theory, based on the claim that "poems are essentially performative" (Folkart, 2007, pp. 59).

A further impact of Meschonnic becomes obvious with regards to the theory of "free verse prosody" developed by authors like Donald Wesling, Richard Cureton, Marjorie Perloff, Burns Cooper, Michael Golston, Jonathan Culler and Derek Attridge. Being influenced by Meschonnic's approach, these theorists discovered rhythmic patterns beyond classical metrics in free verse poetry. On the other hand, leading theorists in "free verse prosody" like Cureton, Attridge, or Wesling became more and more important in current translation theory Boase-Beier (2011); Scott (2015); Underhill (2016); Scott (2018).

Our paper will focus on two theorists within free verse prosody research, both being deeply influenced by Meschonnic: Donald Wesling's theory on "grammetrical ranking", following Meschonnic's call to "go beyond the confines of metrics and linguistics" (Wesling, 1996, pp. 22). And Richard Cureton's concept of Cureton (1992), which is based to Meschonnic's concept of "the speaking subject as discourse, in and by discourse" (Cureton, 1992, pp. 69).

The paper is organized as follows: Section 2 presented the database, the detection of the 'Poetics of Translation' is reviewed in Section 3. Section 4 shows the modeling of the 'Poetics of Translation'. Finally, conclusions and future works are presented in Section 5.

2 Database

*Lyrikline*¹ is an international website and network for experiencing the diversity of contemporary poetry by listening to the sounds and rhythms of international poetry, recited by the authors themselves. Beside, it contains poems both in their original language and in various translations. They are based on a simple rule: Each author has to translate another, which means each partner is asked to contribute the same amount of translations into the project as they contribute new poems to the site: new ones, old ones, from the partner's language into others, or translations into their own language. This explains the huge amount of all in all 17,000 translations from all in all 82 languages, spread over more than 1,300 poets and nearly 12,000 poems. Figure 1 shows the distribution of translation data between four languages (German, English, French, and Dutch).

	Total	German SP	English SP	French SP	Dutch SP
Total		2491	1406	708	524
German TP	6922		900	477	420
English TP	2861	691		38	208
French TP	1262	362	40		144
Dutch TP	421	227	14	9	

Fig. 1: The distribution of translation data for four languages in *lyrikline*.

The aforementioned spoken form of each source poem offers our project further insights into those “paralinguistic features” like duration, pausing, tempo, loudness, tone, and intonation. Clive Scott identified as the main challenge in postmodern poetry translations (Scott, 2015, pp. 46). We already analyzed those features and patterns within the *lyrikline* corpus, who are not based on the metric foot (Iamb, Trochee or Anapest) but rather on free verse prosodic patterns imitating everyday language or musical styles such as jazz, bebop or hip hop Meyer-Sickendiek et al (2018). We sorted such patterns according to a fluency-disfluency scale, starting with the “parlando” developed by Gottfried Benn, and ending with the “lettristic decomposition” developed by Dadaist Isidore Isou. This sorting according to the fluency-disfluency criterion will orientate the translation mapping system described in this paper.

¹ <http://www.lyrikline.org>

3 Detecting the ‘Poetics of Translation’

Meschonnic's “poétique du traduire” Meschonnic (1999) as well as his “critique du rythme” Meschonnic (1982) are both based on his numerous translations of modern and postmodern poetry. Both these volumes are calling for a new awareness for the rhythm and the “discours” of a source text while translating.

3.1 Translating Poetry with Regards to the Grammetrical Features: The “Grammetrical Ranking”

The first important approach in the free verse prosody is the Donald Wesling's theory on grammetrical ranking, which is directly connected to Meschonnic's concept of going “beyond the confines of metrics and linguistics” (Wesling, 1996, pp. 22).

The idea of grammetrical ranking was developed by Donald Wesling with regards to poets like William Carlos Williams, Charles Olson, Allen Ginsberg, Robert Duncan, David Antin, Charles Bernstein, and Ron Silliman. In order to capture these poets “emphatic version of prosody-as-rhythmic-cognition view” (Wesling, 1996, pp. 43), Wesling coined the term grammetrics, a hybridization of grammar and metrics, based on the key hypothesis is that the interplay of sentence structure and line-structure can be accounted for more economically by simultaneous than by successive analysis. In poetry as a kind of versified language, the singular sentence interacts with verse periods (syllable, foot, part-line, line, rhymed pair or stanza, whole poem), a process for which Wesling finds ‘scissoring’ an apt metaphor: Grammetrics assumes that meter and grammar can be scissored by each other, that the cutting places can be graphed with some precision: “One blade of the shears is the meter, the other is the grammar. When they work against each other, they divide the poem. It is their purpose and necessity to work against each other” (Wesling, 1996, pp. 67).

In Wesling's scheme (see Figure 2), the vertical axis designates the grammatical rank and the horizontal axis the metric rank. Intersections of the two axes are represented by circles in which the axes meet; small circles for small coordinate points, large circles for large ones. By using Wesling's scheme, we can compare the grammetrical structure in a source and a target poem, for instance with regards to the specific use of enjambments (weak enjambment, based on clauses, or strong enjambment, based on groups or words?). The following ex-

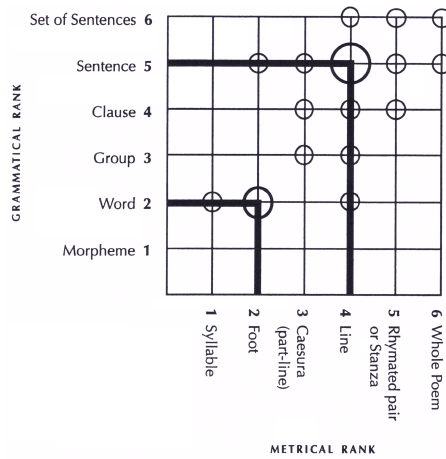


Fig. 2: Grammetrical ranking scheme according to Donald Weslings “The Scissors of Meter” Wesling (1996).

ample (see Table 1) shows how to reproduce the grouping structure used in the original text in translation. In her poem, Kerstin Hensel uses a rhythm coined by Brecht, which Brecht described as “gestic”. Its basic principle is the use of “strong enjambments” occurring for example in line 1 and 2: the transitive verb “to move something away” (*fortrücken*) demands a direct object that appears in the second line: “the table”: “... rückte er//den Tisch fort ...”. In the grammatical ranking, this strong enjambment has a (3) or (2) on the vertical axis and a (4) on the horizontal axis. Hensel, following Brecht, is concerned with alienating the normal rhythm of speech, which becomes clear if one takes a short break at the end of each line: Because the listener or reader expects this direct object, the meaning of the sentence is irritated by the pause in speech.

Dorothy Porter's translation lacks the alienation effect of Brecht's “gestural rhythm”, which becomes obvious when reading a similar pause after each line. The alienating effect is missing, because Porter only uses “weak enjambments” dividing the lines into “kola” i.e. natural pauses in speech without separating transitive verb and object: “he moved the table away” therefore occurs in one line not being separated. In the grammatical ranking, this weak enjambment has a (4) on the vertical axis and a (4) on the horizontal axis. A better translation in this case would be: “when I went to him he//moved the table away and leaned//the bed against the wall ...”

Table 1: Dorothy Porters translation of Kerstion Hensel's poem "Als ich bei ihm war" from German to English.

Kerstin Hensel	Dorothy Porter
ALS ICH BEI IHM WAR RÜCKTE ER Den Tisch fort und das Bett Lehnte er steil an die Wand, und er legte Mich zwischen sich und dem was da anfang Girlanden von Träumen	WHEN I WENT TO HIS PLACE he moved the table away and leaned the bed against the wall - but me he laid between himself and the flowering beginning of our dreams.

3.2 Translating Poetry with Regards to the Rhythmical Features: The "Rhythmic Phrasing"

The second important approach in free verse prosody is Richard Cureton's theory on rhythmic phrasing in English verse, which is directly connected to Meschonnic's concept of "the speaking subject as discourse, in and by discourse" (Cureton, 1992, pp. 69).

Cureton has divided the poetic rhythm into three components: meter, grouping, and prolongation (Cureton, 1992, pp. 125). The meter contains the perception of beats in regular patterns, the grouping refers to the linguistic units gathered around a single climax or peak of prominence, quite similar to Weslings ranking. The new Cureton's idea about prolongation refers to the anticipation and overshooting of a goal, the experience of anticipation and arrival.

Rhythmic prolongation is a matter of connected, goal-oriented motion, based on three levels: anticipation (a), arrival (r), and extension (e) (Cureton, 1992, pp. 146f). In the syntax of "The plowman homeward plods his weary way", the syntactic subject "The plowman (w-a)" anticipates the syntactic predicate "plods his weary way (s-r)", while the more freely positioned and connected adverbial "homeward (w-e)" is felt as a linear extension of the subject before the predicate arrives (see Figure 3).

Our second example for poetry translations (see Table 2) will show how to compare this prolongation of a source and a target text with regards to a famous poem written by Rolf Dieter Brinkmann. The most important aspect in the source text becomes obvious with regards to the paratactic sentence in the beginning: The reader expects the verb "hören", which normally occurs in the

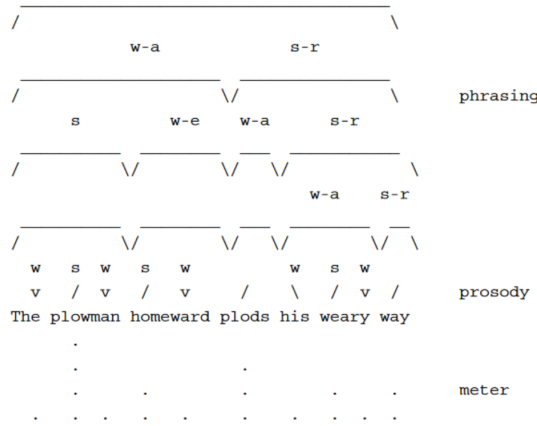


Fig. 3: Rhythmic phrasing schemes according to Richard Cureton's “Rhythmic Phrasing in English Verse” Cureton (2015)

beginning of a sentence, getting somehow breathless when reading the first 9 lines (starting from the headline). This kind of interplay between anticipation (of a verb following the accusative from line 1 and 2) and arrival (of the verb occurring in line 7), interrupted by the enormous extension of the paratactic lines (2 - 7) is exactly what Cureton calls a “prolongation”, which constitutes the rhythmical phrasing of this poem.

It is obvious when reading already the first line of the translation, that Hartmut Schnell failed in keeping this extreme prolongation in his target poem: The verb “to hear” already occurs in the first line, reducing the prolongation from 6 lines to 0. For this simple reason of reducing the prolongation of the source poem, the feeling of getting “breathless” doesn't occur in a similar fashion within the target poem. A better translation in this case would be: “To hear in Cologne, at the end ..., one of those classic black tangos”.

4 Modeling the ‘Poetics of Translation’

The aforementioned two basic elements in free verse prosody - the grammatical ranking and the rhythmic phrasing - are the most important aspects when testing the quality of free verse poetry translations. We consider any poem as a hierarchical arrangement according to the grammatical ranking with regards

Table 2: Hartmut Schnell's translation of Rolf Dieter Brinkmann's poem "Einen jener klassischen" from German to English.

Rolf Dieter Brinkmann	Hartmut Schnell
Einen jener klassischen schwarzen Tangos in Köln, Ende des Monats August, da der Sommer schon	To hear one of those classic black tangos in Cologne, at the end of the month of August, when the summer is
ganz verstaubt ist, kurz nach Laden Schluß aus der offenen Tür einer	already all covered with dust, just after shops close from the open door of a
dunklen Wirtschaft, die einem Griechen gehört, hören, ist beinahe	dark restaurant that belongs to a Greek, is almost
ein Wunder: für einen Moment eine Überraschung, für einen Moment	a miracle: for a minute a surprise, for a minute
Aufatmen, für einen Moment eine Pause in dieser Straße	a fresh breath, for a minute a pause on this street
die niemand liebt und atemlos macht, beim Hindurchgehen. Ich	that no one loves and makes you breathless when you walk through. I
schrieb das schnell auf, bevor der Moment in der verfluchten	jotted it all down, before that minute faded
dunstigen Abgestorbenheit Kölns wieder erlosch.	again in the damn steamy deadness of Cologne.

to the poem's grammar and metrical arrangement. Regarding this hierarchy, we will construct a high-level "map" that represents the poem's poetic properties, by embedding each line of the poem, as well as the poem as a whole, into vector space. We then attempt to visualize these representations in meaningful ways for the translator. An illustrative example is shown in Figure 4 which displays line-by-line representations of poems according to their grammatical ranking. The translation score of each pair of source-target lines (x-axis) is represented as the distance between their feature vectors in a low-dimensional, continuous-valued semantic space, based on the aforementioned features being important for the translation of free verse poetry: grammatical ranking (y-axis) and rhythmic phrasing resp. prolongation (z-axis). Each single line in this feature-based vector space model (VSM) maps the the distance between source and target poem by identifying grammatical unit and the prolongation with regards to each single line, which is represented via the embedding in the vector space.

We will also exploit vectors to augment the following lines (see numbers on the x-axis) within the whole line arrangement of the source poem with the corresponding lines in the target poem, building bilingually-constrained line vectors. These line vectors will be used to provide additional scoring of line pairs, similar to the standard log-linear framework of phrase-based statistical machine translation.

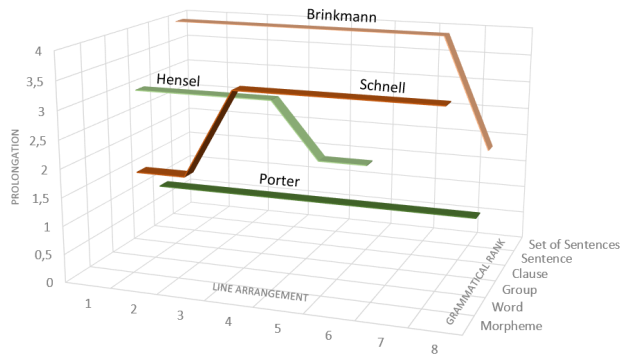


Fig. 4: Simplified illustration of a feature vector space model: x = Line arrangement, y = Grammetrical rank, z = Prolongation

We assume that any poem existing on *lyrikline* can be captured by our representations and classified prosodically. As a result, we expect the model to pinpoint the lack of prolongation in Hartmut Schnell's translation on the very first words (“To hear one of those classic black tangos”), or mismatch of enjambment strength (lack of strong enjambment) in all lines of Dorothy Porter's translation (with the possible exception of the penultimate line).

We construct the poem representation from its hierarchical structure which consists of stanzas, lines, words and whitespace, and finally: characters (or speech and pause in the spoken form). Our representations correspond to the idea of paragraph vectors Mikolov et al (2013); Dai et al (2015) which have been used to improve machine translations Zhang and Lapata (2014); Gao et al (2013). We will hierarchically compose the poem representation from characters and word embeddings to line representations and finally the full poem. We want to optimize embeddings not only based on their distributional semantics, but also their ability to differentiate poetic style using multi-task learning Collobert and Weston (2008). We also intend to investigate further features to build our own representations that are more focused on grammatical and rhythmical

properties. In addition, with the availability of the acoustics of the read source poems in *lyrikline*, we can also partially ground our representations in the verbal form, as well as the results from our prior classification of poetry along the fluency continuum, and other meta-data Meyer-Sickendiek et al (2018).

Poetic style can often be pinpointed to characteristic features of a poem, e.g. the use and kind of enjambments. Again, using multi-task learning, we can direct the model to pay attention to the poem where it is most relevant using supervised training of attention Liu et al (2016). We will base the supervision (a) on the cross-lingual alignments between source and target poems, and (b) on pre-existing annotations (e.g. of enjambment) as well as further annotations to be generated, and (c) pairings of poem translations. We have already demonstrated in previous work Baumann et al (2018) that such prosodic classifications can be traced back to the aforementioned line properties: The grammatical ranking (e.g. strong or weak enjambment), and the rhythmic phrasing (e.g. stressed or unstressed enjambment, prolongation). The attention in the network developed in this work can (a) improve the classification, and (b) orient the focus of the translation or capture the prosodic deviations of each translation. We are planning to use pairs of poems in source and target language in the corpus. This should enable us to learn how a poem's model maps to the target language's vector space for poem models, allowing us to determine whether a given target poem's poetic properties are a "good fit" to the source language poem.

5 Conclusion and Future Works

Our paper described the so-called 'Poetics of Translation' Meschonnic (1999); Folkart (2007); Underhill (2016); Scott (2018) as a framework for automatic poetry translations. By criticizing the traditional domestication of translations, authors like Henri Meschonnic, Barbara Folkart, James W. Underhill or Clive Scott developed a far more individual idea of translation, addressing the foreignness of the source-text as well as the corresponding uniqueness (voice, rhythm, textuality and performativity) of a 'good' translation. We have put a special emphasis on translations within two approaches existing in the theory of free verse poetry: on Donald Wesling's theory on "grammetrical ranking", based on Meschonnic's call to "go beyond the confines of metrics and linguistics" (Wesling, 1996, pp. 22); and on Cureton's rhythmic phrasing in English verse, based on Meschonnic's concept of "the speaking subject as discourse,

in and by discourse” (Cureton, 1992, pp. 69). Showing concrete examples, we demonstrated how to use both these approaches in order to improve translation qualities.

For future works, we are planning to do three further steps: A) To collect and classify translations by analyzing the *lyrikline* corpus, which contains not only a unique amount of recordings of poems read out by the original author, but also a unique amount of all in all 17,000 translations of these poems. B) To develop a recurrent neural network-based analyzer of poetic translations trained by a hands-on feature-classification, leveraging translations performed by translators on the *lyrikline* website. C) To use *lyrikline* translations as training data in order to build an interactive translation mapping system to support the human translator during the work. We want to implement this system as an online translator's workstation on *lyrikline* which will orient each translation to these ‘Poetics of Translation’.

Acknowledgements This work is funded by the Volkswagen Foundation in the announcement ‘Mixed Methods in the humanities? Funding possibilities for the combination and the interaction of qualitative hermeneutical and digital methods’ (funding code 91926 and 93255).

References

- Baumann T, Hussein H, Meyer-Sickendiek B (2018) Analysing the Focus of a Hierarchical Attention Network: the Importance of Enjambments When Classifying Post-modern Poetry. In: Interspeech 2018, Hyderabad, India, pp 2162–2166
- Boase-Beier J (2011) A Critical Introduction to Translation Studies. Bloomsbury Critical Introductions to Linguistics, Bloomsbury Publishing
- Collobert R, Weston J (2008) A Unified Architecture for Natural Language Processing: Deep Neural Networks with Multitask Learning. In: Proceedings of the 25th International Conference on Machine Learning, ACM, New York, NY, USA, ICML '08, pp 160–167
- Cureton R (1992) Rhythmic Phrasing in English Verse. Longman
- Cureton R (2015) Rhythm, Temporality, and Inner Form. *Style* 1(49):78–109
- Dai AM, Olah C, Le QV (2015) Document Embedding with Paragraph Vectors. CoRR abs/1507.07998
- Folkart B (2007) Second Finding: A Poetics of Translation. University of Ottawa Press

- Gao J, He X, tau Yih W, Deng L (2013) Learning Semantic Representations for the Phrase Translation Model. CoRR abs/1312.0482
- Genzel D, Uszkoreit J, Och F (2010) “Poetic” Statistical Machine Translation: Rhyme and Meter. In: EMNLP, pp 158–166
- Ghazvininejad M, Shi X, Priyadarshi J, Knight K (2017) Hafez: an Interactive Poetry Generation System. In: ACL
- Ghazvininejad M, Choi Y, Knight K (2018) Neural Poetry Translation. In: NAACL-HLT
- Greene E, Bodrumlu T, Knight K (2010) Automatic Analysis of Rhythmic Poetry with Applications to Generation and Translation. In: EMNLP, Association for Computational Linguistics, Stroudsburg, PA, USA, EMNLP ’10, pp 524–533
- He J, Zhou M, Jiang L (2012) Generating Chinese Classical Poems with Statistical Machine Translation Models. In: AAAI, pp 1650–1656
- Hopkins J, Kiela D (2017) Automatically Generating Rhythmic Verse with Neural Networks. In: ACL
- Liu Y, Sun C, Lin L, Wang X (2016) Learning natural language inference using bidirectional LSTM model and inner-attention. CoRR abs/1605.09090, URL <http://arxiv.org/abs/1605.09090>, 1605.09090
- Meschonnic H (1982) Critique du rythme. Anthropologie historique du langage. Verdier, Lagrasse
- Meschonnic H (1999) Poétique du traduire. Lagrasse: Verdier
- Meschonnic H (2011) Ethics and Politics of Translating, translated and edited by Pier-Pascale Boulanger. Benjamins Translation Library, John Benjamins Publishing Company
- Meyer-Sickendiek B, Hussein H, Baumann T (2018) Analysis and Classification of Prosodic Styles in Post-modern Spoken Poetry. In: Burghardt M, Müller-Birn C (eds) INF-DH-2018, Gesellschaft für Informatik e.V., Bonn
- Mikolov T, Le QV, Sutskever I (2013) Exploiting Similarities among Languages for Machine Translation. CoRR abs/1309.4168
- Oliveira HG (2012) PoeTryMe: a versatile platform for poetry generation. In: Proceedings of the ECAI 2012 Workshop on Computational Creativity, Concept Invention, and General Intelligence (C3GI at ECAI 2012), Montpellier, France
- Oliveira HG (2017) A Survey on Intelligent Poetry Generation: Languages, Features, Techniques, Reutilisation and Evaluation. In: INLG
- Scott C (2015) The Rhythms of Free Verse and the Rhythms of Translation. *Style* 49(1):46–64

- Scott C (2018) *The Work of Literary Translation*. Cambridge University Press
- Underhill JW (2016) *Voice and Versification in Translating Poems*. University of Ottawa Press
- Wang Q, Luo T, Wang D, Xing C (2016) Chinese Song Iambics Generation with Neural Attention-based Model. CoRR abs/1604.06274, URL <http://arxiv.org/abs/1604.06274>
- Wesling D (1996) *The Scissors of Meter: Grammetrics and Reading*. University of Michigan Press
- Yan R, Jiang H, Lapata M, Lin SD, Lv X, Li X (2013) i, Poet: Automatic Chinese Poetry Composition through a Generative Summarization Framework under Constrained Optimization. In: IJCAI
- Yi X, Li R, Sun M (2016) Generating Chinese Classical Poems with RNN Encoder-Decoder. CoRR abs/1604.01537, URL <http://arxiv.org/abs/1604.01537>
- Zhang X, Lapata M (2014) Chinese poetry generation with recurrent neural networks. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Association for Computational Linguistics, Doha, Qatar, pp 670–680, URL <http://www.aclweb.org/anthology/D14-1074>