

G-101

## 好みに合わせた俳句生成のためのニューラル確率的言語モデルの学習手法の検討 Learning Method of Neural Probabilistic Language Model for Preferable Haiku Generation

加賀 ゆうた 家原 瞭 廣田 敦士 金尻 良介 脇上 幸洋  
Yuta Kaga Ryo Iehara Atsushi Hirota Ryosuke Kanajiri Yukihiro Wakigami  
小西 文昂 松尾 星吾 深田 智 田中 一晶 岡 夏樹  
Bungo Konishi Seigo Matsuo Chie Fukada Kazuaki Tanaka Natsuki Oka

### 1. はじめに

俳句は5・7・5音の合計17音という短い詩であるにも関わらず、「季語」や「切れ」（「や」「かな」「けり」など感動、詠嘆を表す単語で俳句を強制的に切り、読者に想像の余地を与える技法）によって余韻を持たせ、読者に情景を想起させることができる。連歌の発句が独立して俳句となり、松尾芭蕉によってその芸術性が高められ、正岡子規によって近代文芸として確立された。最近では伊藤園 お〜いお茶新俳句大賞において、小中高生の応募が9割以上を締めるなど、若者の関心も高まっている[1]。しかし、そのような機会がなければ実際に俳句を詠む機会は少ない。そのため、自動俳句生成システムを作成すれば、より俳句に触れる機会が増え、俳句文化に貢献できるのではないかと考えた。

近年では、ニューラル確率的言語モデル(Neural Probabilistic Language Model : NPLM)の登場により、統計的言語モデルをニューラルネットワークで扱うことが可能になり、機械翻訳[2][3]や画像キャプション生成[4][5]、自然言語生成の研究が盛んに行われてきた。さらに、現在では、創作課題への取り組みが行われている[6]。

俳句の自動生成については、Wuらは文字レベルのニューラルネットワークによる俳句生成を行った[7]。ニューラルネットワークとして、単純なRNN、Stacked LSTM、Recurrent-Convolutional NN、seqGANを比較し、学習にweb上から収集した俳句を用いた際にはStacked LSTMが、チャットシステムユーザの発話を俳句の形に成形したものを用いた際にはRecurrent-Convolutional NNが優れているという結果を得た。しかし、生成されたモデルの評価はPerplexityによるものに留まっていた。Perplexityとは分岐数を表す評価尺度である。今回の場合では、ある文脈が与えられた時に、次に繋がりうる単語の総数に相当する。言語モデルを仮定すると、その制約により単語の選択肢は減少する。そのため、よい言語モデルを仮定できるとPerplexityは小さくなる。しかし、今回生成するのは俳句という芸術作品であり、その評価を文章としての妥当性だけで測ることは難しく、先行研究においても、人間が読んで良い俳句だと感じるか、という観点からの評価はされなかった。

本研究では、Webから俳句データを収集することを想定している為、Wuらの研究において、学習にweb上から収集した俳句を用いた際に最も評価の高かったStacked LSTMを用いる。そして、人間が良いと感じる俳句、特に「俳句の素人が良いと感じる俳句」が生成できる学習方法を提案する。俳人が詠んだ俳句は芸術性が高く、俳句の素

人には難解なことが多い。一方で、俳句の素人が詠んだ俳句は理解しやすいが、よい俳句と言えるものの数が少ない。そこで、提案手法では、まず俳人が詠んだ多数の良い俳句を用いて学習をさせ、その後俳句の素人が詠んだ少数の良い俳句でモデルを調整する。データを分けずに学習するモデルと比較して、アンケートによる評価が向上するかを検証する。

### 2. ニューラル確率的言語モデル

ニューラル確率的言語モデル(Neural Probabilistic Language Model : NPLM)とは統計的言語モデルをニューラルネットワークで構成したものである。統計的言語モデルとは、ある単語列が与えられた時に、その次に来る単語を予測するモデルであり、単語列 $w_1, w_2, \dots, w_i$ に対して確率 $P(w_{i+1}|w_1, w_2, \dots, w_i)$ を計算する。Bag-of-Wordsとは異なり語順を考慮するのが特徴である。しかし、語順を考慮すると、単語列 $w_1, w_2, \dots, w_i$ のパターン数が膨大となり、そのほとんどが観測されない問題が起きる。また、単語の表現には一般にOne-hotベクトルが用いられるため、学習していない単語に対しては、他の単語の学習結果を全く活かさない。それに対し、ニューラル確率的言語モデルでは単語を分散表現と呼ばれる低次元のベクトルで表現することでその問題を解決する。似た文脈に現れる単語からは似た分散表現が得られるため、ある単語Aについて十分な学習が行われていれば、その単語Aと似た文脈に現れる単語Bについての学習が不十分でも単語Aによる学習結果が単語Bに対して利用できる。

### 3. 手法

#### 3.1. ニューラルネットワーク

実験で使用したニューラルネットワークの構成を図1に示す。入力層への入力には、単語を表すOne-hotベクトルである。しかし、One-hotベクトル表現はスパースな構造であり、そのまま用いては学習に膨大なデータを必要とする。そのため、単語をより小さい次元のベクトルで表現するために、LSTM層の前に一段全結合層を実装している。これによってある単語の学習結果が、その単語と分散表現が似た単語が入力された際にも利用できるようになる。ニューラルネットワークの出力は、 $i$ 単語目までを入力した上でのソフトマックス関数による $i+1$ 単語目の確率分布 $P(w_{i+1}|w_1, w_2, \dots, w_i)$ である。本実験では $i+1$ 単語目として確率が最大のものを選択している。

また、作成したニューラルネットワークのハイパーパラメータを表1に、活性化関数を表2に示す。



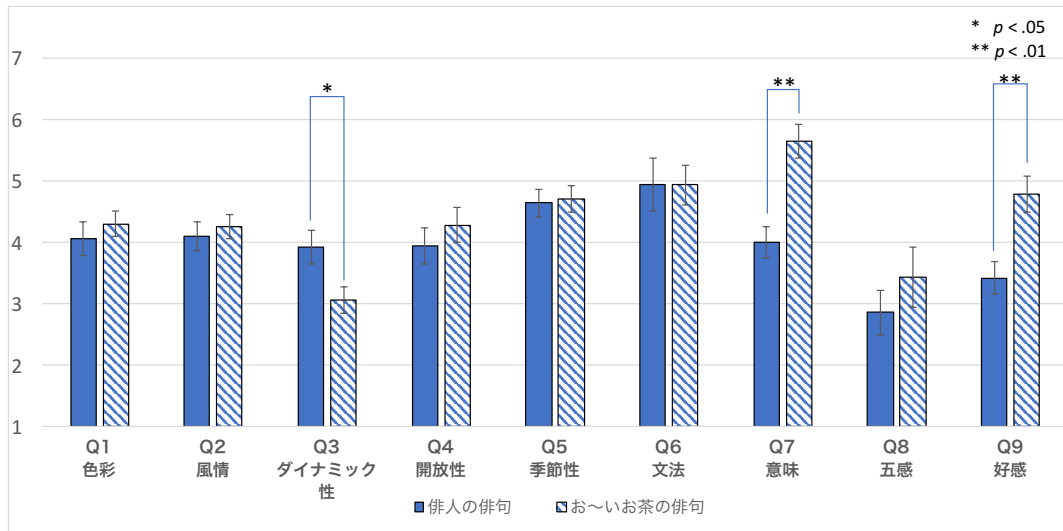


図2: 予備調査結果: お〜いお茶の俳句と俳人の俳句の比較

#### 4.2. 俳句の生成

俳句の生成は以下の手順で行なった。また、今回は「季語」や「切れ字」を挿入する仕組みを作成していないので、生成される俳句は「季語」「切れ字」の無い川柳のようなものとなる。

1. 辞書に載っている全単語を先頭単語とし、3章で述べた2つの学習方法で学習した俳句生成器を用いて俳句を生成する。
2. 生成した俳句と学習データとを比較し、一致率が0.6以上となったものを削除する。
3. 5・7・5音となっていないもの、上五、中七、下五の先頭に助詞・助動詞がきているものを除去する。

一致率の計算方法は生成した俳句に含まれる単語のうち、学習データの個々の俳句に含まれる単語の割合を計算し、その割合が最も高かった俳句における割合をその俳句の一致率とした。なお、閾値は人が見て盗作と感じられるものとそうでないものが分かれるように試行錯誤により選定した。

#### 4.3. ニューラルネットワークの学習

本実験では、生成するものが俳句という創作物であるため、ニューラルネットワークの精度があがるということは俳句の盗作が行われるということの意味する。そのため **early-stopping** のための指標として、完全コピーの個数、4.2節の手順3終了後の俳句の個数、の2つの基準を利用した。

学習を打ち切る段階は、創作物を生成する関係上、完全コピーがそうでないものよりも多く生成されるのは不適切と考え、完全コピーの個数が手順3終了後の俳句の個数を超えない範囲のうち、手順3終了後の俳句の個数が最大となる段階とした。図3～図5に各条件における完全コピー数、手順3終了後の俳句数を示す。

混合条件での学習及び分割逐次条件における前半の学習（俳人の俳句のみでの学習）では2000エポックの時点で完全コピー数が急激に上昇したため、それ以前の1500エポックの時点でのモデルを最良のモデルとした。

分割逐次条件における後半の学習（俳句の素人の俳句のみでの学習）では、学習に用いているデータ数が851個と少ないため、調べた範囲では完全コピー数が手順3終了後の俳句数を上回ることにはなかった。そのため、今回調べた範囲のなかで手順3終了後の俳句数が最大となった225エポックの時点でのモデルを最良のモデルとした。

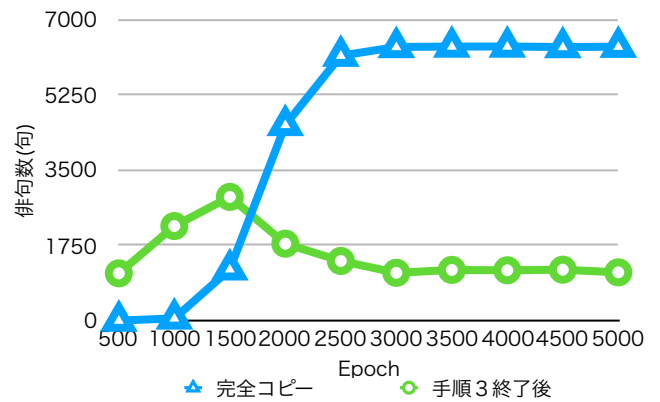


図3: 混合条件の学習曲線

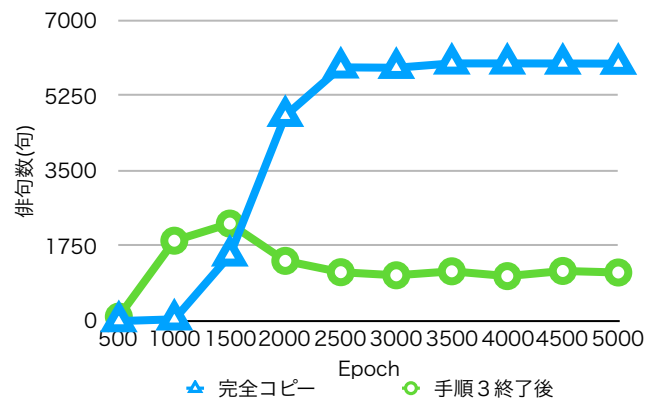


図4: 分割逐次条件における俳人の俳句のみでの学習曲線

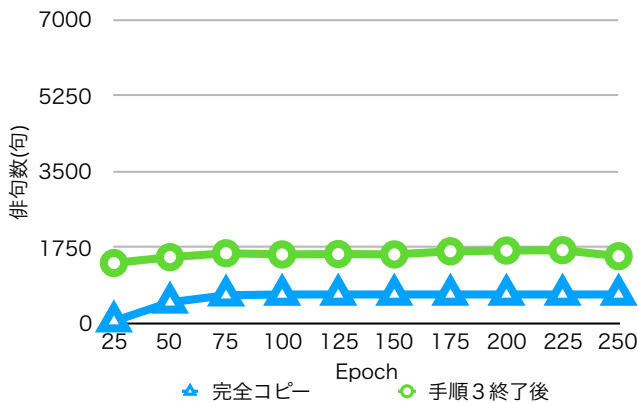


図5：分割逐次条件における俳句の素人の俳句のみでの学習曲線

#### 4.4. 結果

図6に大学生10人を対象に行なった評価実験の結果を示す。混合条件が分割逐次条件よりも有意に季節感があると感じられた( $df=9, t=4.26, p<.01$ )。一方で、分割逐次条件が混合条件よりも有意に俳句の意味が理解され( $df=9, t=-3.49, p<.01$ )、また、有意に俳句が好きだと感じられた( $df=9, t=-4.12, p<.01$ )。いずれの項目も1%水準で有意差が見られた。

生成した俳句に対してアンケート評価を行なった結果と人間が作った俳句に対してアンケート評価を行なった結果に対し、分散分析を行なった結果を図7に示す。俳句に関する印象については、先ほど述べたようにQ5：季節感のない感じ-季節感のある感じで混合条件と分割逐次条件の間に有意差があるのみで、そのほかの設問項目では有意差は見られなかった。一方、Q7：俳句の意味は理解できるについてはお〜いお茶の俳句とその他の3つの俳句の間に有意差が見られた。また、俳人の俳句と混合条件間にも有意差が見られたが、俳人の俳句と混合条件間には有意差が見られなかった。

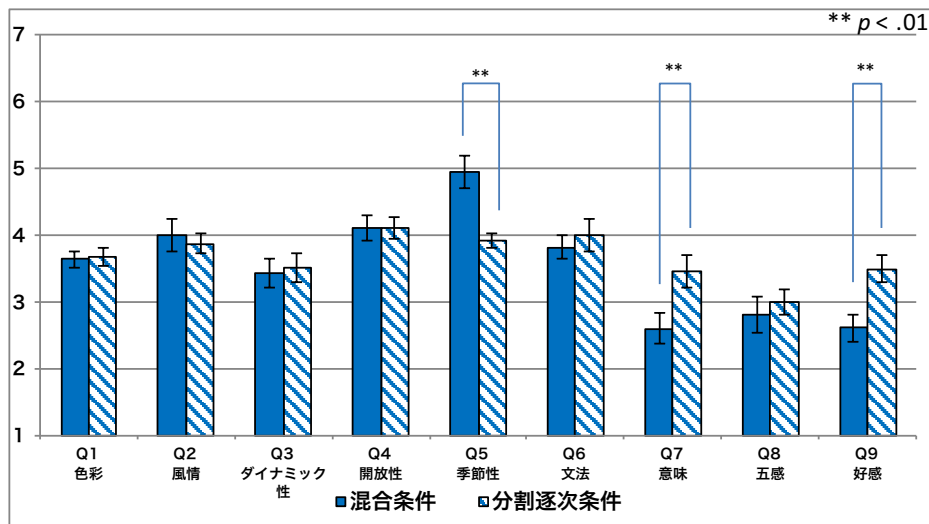


図6：評価実験結果：混合条件と分割逐次条件間の俳句の印象比較

#### 5. 考察

評価実験において、Q7：俳句の意味は理解できる、Q9：この俳句は好きだと思う、の2項目で分割逐次条件の方が混合条件よりも評価点が有意に高かった。これは、分割逐次条件の方がより俳句の素人が良いと感じる俳句を生成できたことを意味する。また、予備実験でもこの2項目では有意差が見られたことから、混合条件はより俳人の俳句らしい俳句が生成され、分割逐次条件ではよりお〜いお茶の俳句らしい俳句が生成されると考えられる。

Q5：季節感のない感じ-季節感のある感じで混合条件の方が分割逐次条件よりも評価点が有意に高かった理由について考える。俳人の俳句とお〜いお茶の俳句の印象を調査した予備実験において、Q5：季節感のない感じ-季節感のある感じでの評価理由についての自由記述を見ると、俳人の詠んだ俳句については、「耐震の壁に書初貼られけり」という俳句に対し「書初という単語で季節感を感じた」というような一つの単語から季節感を抱いている回答が俳句5個中4個で見られた。一方でお〜いお茶の俳句については、「妹の下駄をそろえる月あかり」という俳句に対して「夏祭りのあとの家の玄関を思い浮かべた。」のように、いくつかの単語からある情景を思い浮かべて、そ

れによって季節感を感じている回答が俳句5個中3個で見られた。混合条件ではデータセット中の学習データに、俳人の俳句のデータ量が多かったため、俳人の俳句に偏った学習をされると考えられる。その結果単独の単語で季節感を表現できる単語が生成されやすくなったのではないかと考えられる。一方で分割逐次条件では、後半の学習データであるお〜いお茶の俳句に学習が引っ張られるため、単独の単語で季節感を表現できる単語が生成されにくくなったと考えられる。

Q7、Q9の2項目では混合条件と分割逐次条件間で有意差は見られたが、その平均値は高い評価を得た分割逐次条件でもスコアが4.0以下であった。この理由としては、俳句の生成手順で行なった、盗作の除去にあると考えられる。ニューラルネットワークは教師データを再現できるように学習する。そのため、「俳句らしいフレーズ」を再現できるようになっていくのだが、再現しすぎてしまうと創作物として不適切なものになってしまう。そのため、俳句らしいフレーズが含まれていないものが結果的に残り、人間が読んだときに不自然に感じてしまったのではないかと考えられる。

図7より、俳句の意味の理解という観点において、分割

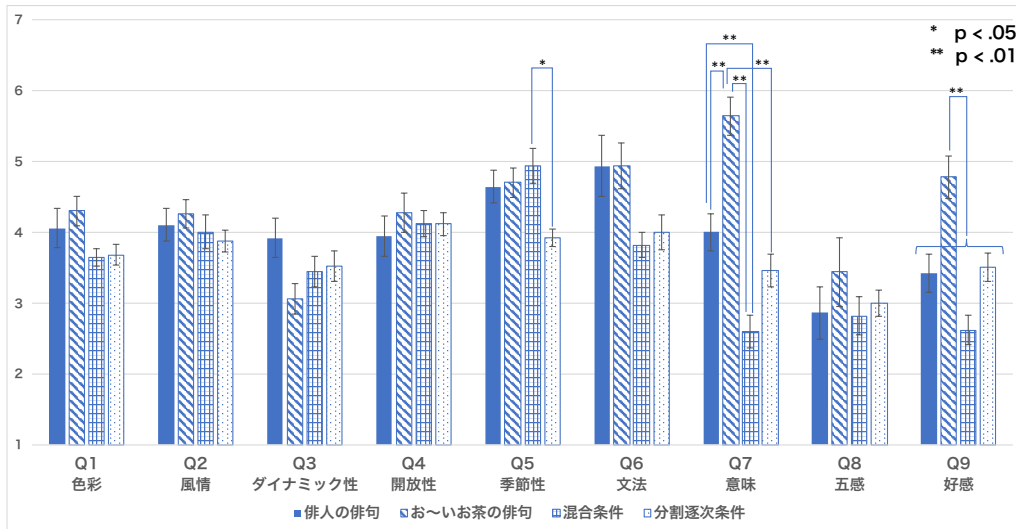


図 7：4 種類の俳句に対するアンケート結果

逐次条件はお〜いお茶の俳句には敵わないものの、俳人の俳句と差があるとは言いきれない。つまり、俳人の俳句と張り合えるだけの俳句が作れているのではないかと考える。

今回の実験では、俳句はランダムに選択された5個ずつのみ使用しているため、異なる俳句を選んだ場合についての議論ができていない。今後は、被験者ごとに提示する俳句を変えるなどして、選択する俳句が変わっても結果が変わらないかどうか検証する必要があると考えられる。

また、学習データの個数の比率がどの程度までであれば、少数のデータセットの特徴を反映させることができるのかも検証していきたい。

## 6. 結論

本研究では俳句の素人が良いと感じる俳句を生成することを目指し、初めに、俳人が詠んだ俳句で学習を行い、その後俳句の素人の詠んだ俳句のみで学習を行う手法を提案した。この手法は、俳人が詠んだ俳句及び俳句の素人が詠んだ俳句をひとまとめにしたものをデータセットとして学習を行う手法と比較して、意味が理解されやすく、好まれやすい俳句が生成できることを確認した。

現在は、あるタイミングで完全にデータセットを切り替えて学習を行っているが、今後はその切り替えを線型的に変化させるなど工夫することで、データセットを完全に切り替えるよりも誤差の急激な変動が小さくなり、より自然で好まれやすい俳句が生成できるのではないかと考えている。

## 7. 参考文献

- [1]第 28 回 伊藤園 お〜いお茶新俳句大賞  
「<https://www.itoen.co.jp/new-haiku/>」  
最終アクセス日：2017 年 7 月 21 日
- [2]Ilya Sutskever, Oriol Vinyals, and Quoc V. Le : Sequence to sequence learning with neural networks, Advances in Neural Information Processing Systems, pp.3104-3112, (2014).
- [3]Yonghui Wu, Mike Schuster, Zhifeng Chen et al. : Google's neural machine translation system: Bridging the gap between human and machine translation, arXiv preprint arXiv:1609.08144, (2016).
- [4]Andrej Karpathy, and Li Fei-Fei : Deep visual-semantic alignments for generating image descriptions, Proceedings of the

- IEEE Conference on Computer Vision and Pattern Recognition, pp.3128-3137, (2015).
- [5]Oriol Vinyals, Alexander Toshev, Samy Bengio et al. : Show and tell: A neural image caption generator, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.3156-3164, (2015).
- [6]Samuel R. Bowman, Luke Vilnis, Oriol Vinyals et al. : Generating sentences from a continuous space, arXiv preprint arXiv:1511.06349, (2015).
- [7]Xianchao Wu, Momo Klyen, Kazushige Ito et al. : Haiku generation using deep neural networks, 言語処理学会 第 23 回 年次大会 発表論文集, (2017).
- [8]積田 洋, 竹内 政裕, 鈴木 弘樹 : 俳句から連想する心象風景の構成と心理的評価の研究, 日本建築学会計画系論文集 第 76 卷 第 669 号, pp.2093-2099, (2011).
- [9]鈴木 雅実 : 感性コミュニケーションメディアとしての俳句, 人工知能学会誌 Vol. 21 No.2, pp.189-194, (2006).