俳句の韻律の境界と構文構造上の境界は一致するか

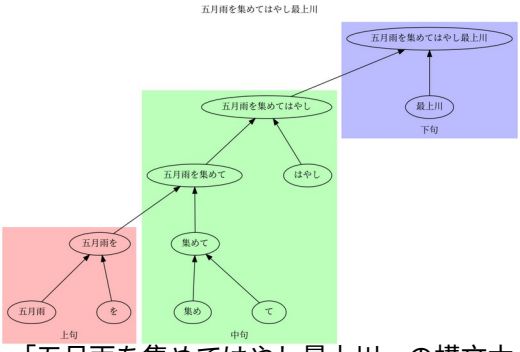
23B13000 松浪大輔
東京工業大学工学院

1．はじめに

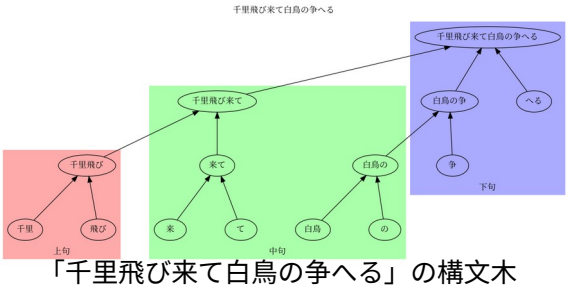
俳句構文構造上について言及した先行研究は通常の日本語文法からの逸脱やそれにより産まれる効果に着目したものが多く、俳句の韻律の境界と構文構造上の境界が一致するかという点が不明であった。よって本レポートでは、「境界が一致する」ということの定式化と、それにそった調査と考察を行った。

2．方法

まず、「境界が一致する」ことをグラフ理論による定式化を行う。ここで、「五月雨を集めてはやし最上川」と「千里飛び来て白鳥の争へる」の二句を例とする。



「五月雨を集めてはやし最上川」の構文木



「千里飛び来て白鳥の争へる」の構文木

以上が二句の構文木である。ここで、各語は上句・中句・下句のいずれかに属するものとし、葉要素以外はＸバー理論における主要部の属する韻律上の区分に属するものとする。そして、「境界が一致する」ことを、すべての韻律上の区分について、それに属するすべての要素の最近共通祖先 (LCA) がその区分に属することと定義する。

この定義に従うと、「五月雨を集めてはやし最上川」は境界が一致し、「千里飛び来て白鳥の争へる」は一致しない。
この定義に従い、俳句のデータセット https://huggingface.co/datasets/p1atdev/modern_haiku より破調の句を除き 500 句を無作為に抽出して、これらについて境界が一致するかどうかを判定する。

3．結果

調査を行った 500 句のうち 237 句は境界が一致していた。これは 47% にあたる。

4．考察

これらの結果から、境界が一致していることは俳句であることの必要条件ではないということがわかった。

5．おわりに

本レポートでは hugging face 上のデータセットを用いたが、本データセットは江戸以前の俳句を収録していないことから、江戸時代の俳句を考慮することができていないという問題がある。また、 47% という値が統計的な見地からどのような評価をしうるのかという点に関しても考察することができなかった。
なお、韻律の分割は手動で、構文解析には GiNZA と spiCY を利用した。

文献：
新田 義彦『俳句における補文構造とその省略の種々相』経済集志 = The Nihon University economic review / 日本大学経済学部 編 83 (1), 13-26, 2013-04
渡辺　明『生成文法』 2009 年　東京大学出版会