The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

スタイルの類似性を捉えた単語ベクトルの教師なし学習

Learning Style-sensitive Word Vector via Unsupervised-manner

赤間怜奈 *1 Reina Akama　渡邉研斗 *1 Kento Watanabe　横井祥 *1*2 Sho Yokoi　小林颯介 *3 Sosuke Kobayashi　乾健太郎 *1*2 Kentaro Inui

*1東北大学 大学院情報科学研究科 Graduate School of Information Sciences, Tohoku University
*2理研 AIP RIKEN AIP
*3株式会社 Preferred Networks Preferred Networks, Inc.

This paper is the first study aiming at capturing stylistic similarity in an unsupervised manner. We construct a novel style-sensitive word vector predicting the target word for giving nearby and wider contexts under the assumption that the style of all the words in an utterance is consistent. We also introduce an evaluation dataset with human judgments on the stylistic similarity between word pairs. Experimental results illustrate capturing the stylistic similarity significantly.

1. はじめに

自然言語において，**何を伝えるか（意味）**だけに限らず，**どう表現するか（スタイル）**の側面も考慮することは重要であり，ことばのスタイルをモデル化することは，文書分類や文生成などの自然言語処理に有用である [1, 2, 3].

このような背景のため，近年では，スタイルのある側面に部分的な定義（たとえば，性別や敬意表現など）を与えることで，スタイルをモデル化する研究が盛んに行われている [4, 5, 6, 7, 8]. しかし，スタイルは極めて主観的な性質を持つため，スタイルそのものを明確に定義することは難しい [9]. そのため，定義されたスタイル-ラベルの分類問題などの教師ありタスクに定式化することは困難かつ不適切である.

本研究の貢献を以下に示す：(1) スタイルを定義せずに，スタイルの類似性を捉えた単語ベクトル空間を教師なし学習する. ここで我々が期待する単語ベクトル空間とは，(「意味」は近くとも)「スタイル」が大きく異なる “俺” と “私” は遠くに配置され，(「意味」は異なっているとしても)「スタイル」が似ている “俺” と “だぜ” が近くに配置されるような空間である（図 1 左）. (2) スタイルの類似性を包括的に定量評価する新しい手法を提案し，そのための評価用データセットを作成する. 本研究はスタイルの類似性を定量的に評価するための評価データセットを作成した最初の研究である. (3) 提案手法により獲得した単語ベクトルが，スタイルの類似性を捉えていることを定量的および定性的に示す.

2. スタイルの類似性を捉えた単語ベクトル

2.1 本論文で用いる記号の定義

本論文では，発話を単語列 w_1,w_2,...w_I として表す. ここで I は発話に含まれる単語数である. たとえば w_i は発話内の i 番目の単語を表す. とくに注目するターゲット単語を w_t と表し，w_{t+1} はターゲット単語の次の単語を表す. また，$\boldsymbol{v}_w$ と $\tilde{\boldsymbol{v}}_w$ はそれぞれターゲット単語ベクトルと文脈単語ベクトルを表す. ターゲット単語と文脈単語との距離を d，文脈窓のサイズを δ とする.

連絡先: 東北大学 大学院情報科学研究科
〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6-05
E-mail: reina.a@ecei.tohoku.ac.jp

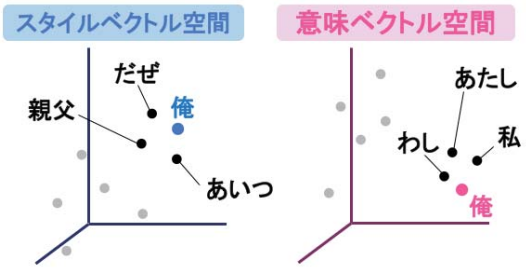

図 1　一般的な単語ベクトル空間（右）では，意味的に類似した単語が近い空間に配置される. 一方で，本研究で期待するスタイルを捉えたベクトル空間（左）とは，たとえば（意味は異なるとしても）スタイルが類似する “俺” と “だぜ” が近くに配置され，(意味は近くとも) スタイルが大きく異なる “俺” と “私” が遠くに配置されるような空間である.

2.2 ベースラインモデル

まず，Mikolov らが提案した Countinuous Bag-of-Words (CBOW) [10] の概要を説明する. 彼らが「vector representations that capture a large number of precise *syntactic* and *semantic* word relationships」と述べているように [11], CBOW は単語の統語情報および語義情報等の「意味」の類似性を捉えた単語ベクトルを学習する確率モデルである. CBOW では，近傍の文脈単語集合 $C_{w_t}^{near} = \{w_{t\pm d} | \ 1 \leq d \leq \delta\}$ (δ は文脈窓幅である) を入力とした時のターゲット単語 w_t の予測確率を最大化することにより，単語ベクトル $\boldsymbol{v}_{w_t}$ を学習する：

$$P(w_t|C_{w_t}^{near}) \propto \exp(\boldsymbol{v}_{w_t} \cdot \tilde{\boldsymbol{v}}_{C_t^{near}}), \qquad (1)$$

ここで，$\tilde{\boldsymbol{v}}_{C_t^{near}} := \frac{1}{|C_{w_t}^{near}|} \sum_{w \in C_{w_t}^{near}} \tilde{\boldsymbol{v}}_w$ は，全文脈単語ベクトルの平均ベクトルを表す. 本研究では，この CBOW をベースラインモデル *CBOW-near-ctx* と呼ぶ.

2.3 提案モデル 1：発話内のスタイル一貫性に基づくモデリング

本研究では，単語のスタイルの類似性を捉えた単語ベクトルを学習したい. しかし，*CBOW-near-ctx* のように近傍の文脈単語集合からターゲット単語を予測するだけでは，単語の「意味」の類似性を捉えた単語ベクトルを学習できても，スタイルの類似性を捉えた単語ベクトルは学習できないと考えられる. スタイルの類似性を捉えるための手がかりとして，性別属性 [12]

や人物情報 [13] などが有用と考えられるが，同じ「女性」という属性であっても，全ての女性が必ずしも同一のスタイルで発話するとは考えにくい．さらに，同一人物であっても，話し相手や状況に応じて複数のスタイルを使い分けることは珍しくない．したがって，性別情報や人物情報ではスタイルの類似性を捉えることは難しいと考えられる．

そこで本研究では，性別属性や人物情報を使わずにスタイルの類似性を捉える手がかりとして，「属性や人物情報に関わらず，ある 1 つの発話内の全単語のスタイルは一貫している」という直観的な仮定をたてた．この仮定に基づき，発話内の全ての文脈単語 $C_{w_t}^{all} = \{w_{t\pm d} | \ 1 \leq d \leq I\}$ からターゲット単語を予測するモデル *CBOW-all-ctx* を学習することで，スタイルの類似性を捉えた単語ベクトルを構築する．これは，ベースライン *CBOW-near-ctx* の文脈窓幅 δ を発話長 I に拡張することに等しい．この拡張モデルでは，意味の類似性に加えてスタイルの類似性も捉えた単語ベクトルが学習されると期待される．

2.4 提案モデル 2：スタイル情報と意味情報の分離

前節では意味の類似性とスタイルの類似性を捉えるモデル *CBOW-all-ctx* を提案した．本研究では *CBOW-all-ctx* を拡張して，純粋にスタイルの類似性のみを捉えた単語ベクトルの獲得を試みる．

2.2 節では，近傍の文脈単語からターゲット単語を予測することで，意味の類似性を捉えた単語ベクトルが学習可能であると述べた．したがって，近傍の文脈単語を無視すれば，意味の類似性を捉えない単語ベクトルが学習されると考えられる．そこで，スタイルの類似性のみを捉えるシンプルな工夫として，*CBOW-all-ctx* の学習に用いた文脈語から，近傍 $C_{w_t}^{near} = \{w_{t\pm d} | \ 1 \leq d \leq \delta\}$ を除外する．この工夫により，スタイル情報だけを捉えた単語ベクトルが学習されると期待される．このモデルを *CBOW-dist-ctx* と呼ぶ．

さらに本研究では，単語ベクトル $\boldsymbol{v_t}$ を (1) スタイルの類似性のみを捉える単語ベクトル $\boldsymbol{x_t}$ と (2) 意味の類似性のみを捉える単語ベクトル $\boldsymbol{y_t}$ に分離しながら同時に学習する拡張モデル *CBOW-sep-ctx* を提案する（図 2）：

$$\boldsymbol{v}_{w_t} = \boldsymbol{x}_{w_t} \oplus \boldsymbol{y}_{w_t} \tag{2}$$

$$\tilde{\boldsymbol{v}}_c = \tilde{\boldsymbol{x}}_c \oplus \tilde{\boldsymbol{y}}_c \tag{3}$$

ここで，$\oplus$ はベクトルの連結を表す．また，$\boldsymbol{x_t}$ をスタイルベクトル，$\boldsymbol{y_t}$ を意味ベクトルと呼ぶ．単語ベクトル $\boldsymbol{v_t}$ と同様に，文脈単語ベクトル $\tilde{\boldsymbol{v}}_c$ もスタイルベクトルと意味ベクトルに対応するベクトル $\tilde{\boldsymbol{x}}_c$，$\tilde{\boldsymbol{y}}_c$ に分割される．

本研究では，スタイルの類似性と意味の類似性を明示的に分離して学習するために，ターゲット単語から文脈単語までの距離に応じてパラメータを更新するベクトルを切り替える．具体的には，ターゲット単語の近傍の文脈単語集合 $C_{w_t}^{near}$ からは意味の類似性とスタイルの類似性の両方を捉えることができると考えられるため，スタイルベクトル $\boldsymbol{x}_{w_t}$ と意味ベクトル $\boldsymbol{y}_{w_t}$ の両方のパラメータを更新する．また，ターゲット単語から一定の距離 δ 以上離れている文脈単語集合 $C_{w_t}^{dist}$ からはスタイルの類似性のみを捉えることができると考えられるため，スタイルベクトル $\boldsymbol{x}_{w_t}$ のパラメータのみを更新する．以上の更新手順をまとめると，予測確率 $P(w_t|C_{w_t})$ は以下のように定式化できる：

$$P(w_t|C_{w_t}) \propto \begin{cases} \exp(\boldsymbol{v}_{w_t} \cdot \tilde{\boldsymbol{v}}_{C_{w_t}^{near}}) \ (1 < d \leq \delta) \\ \exp(\boldsymbol{x}_{w_t} \cdot \tilde{\boldsymbol{x}}_{C_{w_t}^{dist}}) \ (\delta < d \leq I) \end{cases} \tag{4}$$

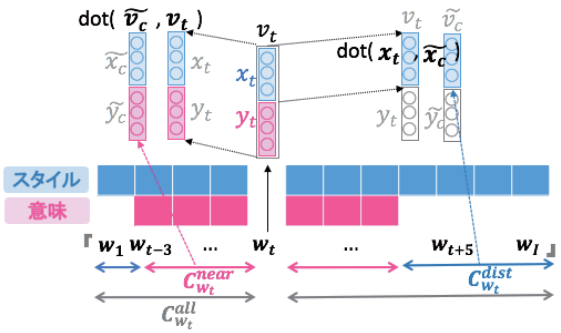

図 2　*CBOW-sep-ctx* の概念図．ターゲット単語の近傍で共起する単語（$C_{w_t}^{near}$）はスタイル（青）および意味（赤）の両方の類似性に有効かつ発話内で共起する単語（$C_{w_t}^{all}$）はスタイルの類似性に有効という考えに基づく．

3. 実験

提案した単語ベクトルがスタイルおよび意味（つまり統語，語義）の類似性を捉えているか評価する．

データセット

本研究では，学習コーパスとして，ショートストーリー (SS) を用いる．SS とは，Web に投稿された漫画やアニメの二次創作作品であり，かぎ括弧で囲まれた登場人物の発話が記載されている．我々は，かぎ括弧で囲まれた約 30M 発話を収集した．前処理として，収集した発話に対して単語単位の分かち書きや品詞情報の付与等 *1 を施した後，スタイルを捉えやすくする工夫として “です + わ” や “です + ぜ” のように文末の機能語の連結処理をおこなった [12]．作成したコーパスの 99% を訓練データ，1% をテストデータとして用いた．

パラメータ設定

学習する全ての単語ベクトルの次元数を 300 次元に統一した．例外として，提案モデル *CBOW-sep-ctx* では，分割されたスタイルベクトル $\boldsymbol{x}_{w_t}$ および意味ベクトル $\boldsymbol{y}_{w_t}$ の次元としてそれぞれ 300 次元を用意し，合計 600 次元の単語ベクトル $\boldsymbol{v}_{w_t}$ を学習した．窓幅は $\delta = 5$ を用いた．また *CBOW-sep-ctx* における学習率は，$\{ \boldsymbol{v}_{C_{w_t}^{near}}, \boldsymbol{x}_{C_{w_t}^{dist}}, \tilde{\boldsymbol{v}}_{C_{w_t}^{near}}, \tilde{\boldsymbol{x}}_{C_{w_t}^{dist}} \}$ の各部分で学習率と更新回数の期待値との積が一定となるよう，部分毎に調整をおこなった． 学習にはコーパス内での出現頻度が 5 以上の単語を用い，結果として語彙サイズは約 10 万となった．学習の高速化のため，負例サンプリングによる近似をおこなった．コーパス中の単語の各出現について計 10 回ずつ学習した後，学習を停止した．

3.1 スタイルの類似性に関する評価

提案モデルがスタイルの類似性を捉えていることを検証するために，モデルが捉えた単語の類似性が人間の感じる単語のスタイルの類似性とどの程度相関があるかを定量的に確かめる．

我々が知る限り，スタイルの類似性を定量的に評価するための尺度およびデータセットは存在しない．そこで我々は，単語類似度タスク（similarity task）のための評価データセットの構築手法 [14] を参考にして，以下の 2 ステップで単語ペアに対して人間がスタイルの類似度を付与したデータを新しく構築する．(1) クラウドソーシングによって “俺” や “だぜ” のような何ら

*1　表現の分散による不必要な語彙の増加を防ぐ目的で，動詞の原形化および一部の単語について漢字・ひらがな・カタカナ表現の統一処理をおこなった．

モデル	ρ_{style}	ρ_{sem}	syntax acc @5	syntax acc @10
CBOW-near-ctx	12.1	27.8	86.3	85.2
CBOW-all-ctx	36.6	24.0	85.3	84.1
CBOW-dist-ctx	**56.1**	15.9	59.4	58.8
CBOW-sep-ctx				
スタイルベクトル $\boldsymbol{x}$	**51.3**	**28.9**	68.3	66.2
意味ベクトル $\boldsymbol{y}$	9.6	18.1	**88.0**	**87.0**

表 1　スタイル/語義的/統語的な類似性に関する評価の結果．ρ_{style} および ρ_{sem} は実際の係数 ×100 の値である．

かのスタイルと強く関係しそうな単語を収集する [*2]．具体的には，テストセットからランダムに抽出した 1,000 文の各文を 5 人のアノテータに提示し，4 人以上がスタイルが現れていると判断した単語を選択してもらった．(2) 次に，収集したスタイルに関係する単語の集合から 1,000 の単語ペアをランダムに生成し，各ペアについて 15 人のアノテータに 5 段階（−2：スタイルが異なる〜+2：スタイルが類似している）でスタイルの類似度のスコア付けを指示した．また，単語ペアだけではどのようなスタイルであるかを判定することが難しいため，各単語を含む発話もアノテータに提示した．この発話は，テストセットから無作為に抽出することでひとつの単語ペアにつき 3 種類を用意した．最終的には，アノテータ間での合意が認められた 399 の単語ペアとそれらに対する類似度のスコアが得られた [*3]．本研究はスタイルの類似性を定量的に評価するための評価データセットを作成した最初の研究である．

本研究では，以下の手順で学習された単語ベクトルがスタイルの類似性を捉えているかを定量評価する．(1) 評価データセット内の全単語ペアについて，単語ベクトルを用いてコサイン類似度を計算する．(2) 単語ペアの集合を計算した類似度順に並べ替える．(3) モデルが並べ替えた単語ペアの順序と，人間がスコア付けした単語ペアの順序について，Spearman の相関係数 ρ_{style} を計算する．

結果を表 1 の左部に示す．全ての提案モデル *CBOW-{all,dist}-ctx* 及び *CBOW-sep-ctx* のスタイルベクトル $\boldsymbol{x}$ はベースラインモデル *CBOW-near-ctx* より高い相関を持つことがわかる．特に *CBOW-dist-ctx* とスタイルベクトル $\boldsymbol{x}$ は，それぞれ $\rho_{style} = 56.1$，51.3 と，スタイルの類似性について人間の認識と高い相関を持つ結果となった．これらの結果は，発話内の文脈単語を使用することが，スタイルの類似性を捉えた単語ベクトルの構築に寄与していることを意味する．

3.2　統語・語義の類似性に関する評価

標準的な単語ベクトルである CBOW は，意味（つまり統語および語義）の類似性を捉えることができると検証されている [10]．本研究でも（1）品詞の予測からみる統語的な類似の捕捉，（2）人間との相関からみる語義的な類似の捕捉の 2 つの観点から，提案した単語ベクトルの特性を詳細に分析する．

3.2.1　統語的な類似性

同じ品詞の単語がベクトル空間上の近い位置に配置されているかを計算することで，単語ベクトルが統語的な類似性を捉えているかを評価する．具体的には，あるターゲット単語 w と類似する上位 N 単語 w' の品詞の一致率の平均 syntax acc を以下の式で算出する：

$$\text{syntax acc} = \frac{1}{|V|N} \sum_{w \in V} \sum_{w' \in \mathcal{N}(w)} \mathbb{I}[\text{PoS}(w) = \text{PoS}(w')], \quad (5)$$

ここで $\mathbb{I}[\cdot]$ は指示関数を意味し，$\text{PoS}(w)$ は単語 w の品詞を返す関数を意味する．$\mathcal{N}(w)$ は $\cos(w, w')$ に関する類似度上位 N までに含まれる単語 $\{w'\}$ の集合である．

表 1 の右部に $N = 5, 10$ のときの品詞の平均一致率を示す．どちらの N についても，*CBOW-{near,all}-ctx* および *CBOW-sep-ctx* の意味ベクトル $\boldsymbol{y}$ が統語的な類似性を捉えている結果となった．反対に，*CBOW-dist-ctx* スタイルベクトル $\boldsymbol{x}$ は，統語的な類似性を捉えておらず，これらの結果は，本研究の目的であるスタイルの類似性だけを捉えた単語ベクトルの構築が部分的に達成できたことを意味する．ここで興味深いのは，*CBOW-all-ctx* とスタイルベクトル $\boldsymbol{x}$ は両方とも発話内の全単語を文脈として学習しているにもかかわらず，*CBOW-all-ctx* のほうが統語的な類似性を捉えているということである．これは *CBOW-sep-ctx* の予測確率（式（4））の計算において，近傍の文脈単語 $C_{w_t}^{near}$ から得られる統語的な情報を意味ベクトル $\boldsymbol{y}$ がフィルタリングしたためと推測できる．なお，提案モデル *CBOW-sep-ctx* における $\boldsymbol{x}$ および $\boldsymbol{y}$ がそれぞれ捉える統語的な性質の差異については，3.2.2 節で詳しく分析する．

3.2.2　語義的な類似性

単語ベクトルが語義的類似性を捉えているかを評価する．本研究では，評価データセットとして日本語単語類似性データセット（JWSD） [16] を用いた．JWSD は約 4,000 の単語ペアから形成され，各単語ペアには人間が付与した 0 から 10 の類似度スコアを持つ．我々は，3.1 節のスタイルの類似性評価と同様に，モデルが計算した単語ベクトルのコサイン類似度と人間が付与した類似度スコアとの間の Spearman の相関係数（ρ_{sem}）を算出した．

結果を表 1 の中央部に示す．驚くべきことに，*CBOW-dist-ctx* とスタイルベクトル $\boldsymbol{x}$ はスタイルの類似性評価では高い ρ_{style} であったにもかかわらず，語義の類似性評価ではスタイルベクトル $\boldsymbol{x}$ が高い ρ_{sem} であり，*CBOW-dist-ctx* が低い ρ_{sem} となった．本研究の目的はスタイルの類似性**のみ**を捉えた単語ベクトルの構築であるが，提案したスタイルベクトル $\boldsymbol{x}$ はスタイルの類似性以外の類似性も捉えている結果となった．この原因についての考察は次節で述べるが，我々はスタイルベクトル $\boldsymbol{x}$ が発話のトピック（relatedness）の類似性を捉えたためだと推測した．JWSD には単なる語義の類似だけでなく，あるトピックに関連する単語たちや同義語たちもペアとして含まれているため，トピックの類似性を捉える単語ベクトルでは ρ_{sem} の値は大きくなる．

3.3　スタイルベクトルおよび意味ベクトルの定性分析

提案モデル *CBOW-sep-ctx* によって獲得した 2 種類のベクトルが持つ性質を定性的に分析する．分析対象とするターゲット単語 [*4] について，各ベクトル空間上でコサイン類似度が大きい上位 5 件の単語を表 2 に示す．表 2 左部より，スタイルベク

*2　少なくとも日本語には，役割語 [15] など特にスタイルの差異によく関与する単語がある一方で，「ホワイトボード」などどの人物がどの状況で言っても同じ表現となる，いわばスタイルの差異が関与しない単語がある．そのため，スタイルとは関係ない単語を用いてスタイルの類似性の評価データを作成することは非効率であり，この段階で除外したい．

*3　10 人以上のアノテータのスコアが $\{-2, -1\}$ または $\{+2, +1\}$ に一致している場合を合意があるものとした．

*4　分析対象とする単語は，スタイル類似度評価データセットの作成過程でアノテータ 5 人全員が選択したスタイルに関係する単語を用いた．

ターゲット単語 w_t	スタイルベクトル: $\boldsymbol{x}_{w_t}$	意味 (語義/統語) ベクトル: $\boldsymbol{y}_{w_t}$
俺	おまえ，あいつ，ねーよ，奴，そいつ	僕，あたし，私，おまえ，わし
拙者	でござる，ござる，ござるよ，ないでござる，たでござる	僕，俺，吾輩，私，あたし
かしら	わね, ないわね, わ，だわ，かしらね	かな，でしょうか，かしらね，かなぁ，でしょう
サンタ	サンタクロース，トナカイ，クリスマス，ソリ，プレゼント	お客，プロデューサー，メイド，由比ヶ浜，鹿目

表 2　スタイルベクトル空間および意味ベクトル空間における分析対象の単語とコサイン類似度が大きい上位 5 単語.

トル $\boldsymbol{x}$ 空間では，“俺” に対しては “おまえ” や “ねーよ（文末）” といった男性らしく粗野な単語との類似度が高い結果となった．また “拙者” に対しては “ござる” という忍者や侍らしい古典的な言葉遣いの単語が近くに配置されていることがわかる．このように，スタイルベクトル $\boldsymbol{x}$ はスタイルの類似性を捉えていることが観察された．一方，表 2 の右部より，意味ベクトル空間 $\boldsymbol{y}$ では，“俺” や “拙者” に対しては一人称代名詞である “僕” や “私” が近くに配置されていることがわかる．なお，紙面の都合により省略するが “俺” と類似する上位 15 位の単語全てが人称代名詞であった．これらの結果から，提案した *CBOW-sep-ctx* は (1) スタイルの類似性を捉える単語ベクトルと (2) 統語・語義の類似性を捉える単語ベクトルに分離しながら単語ベクトル同時に学習していることが，定性的に確かめられた．

面白いことに，スタイルベクトル $\boldsymbol{x}$ において単語 “サンタ” は “トナカイ” や “クリスマス” 等の同じ**トピック**の単語との類似度が高いことが観測された（表 2，4 行目）．本研究では，スタイルの類似性と捉えるために「発話内の全単語のスタイルは一貫している」という仮定を立てたが，発話内ではスタイルだけでなくトピックも一貫していることが多い．例えば “サンタ” と “トナカイ” はクリスマスのトピックであり，これらはしばしば 1 つの発話内で共起しやすい．したがって，発話内で共起するの全ての単語の情報を用いたスタイルベクトル $\boldsymbol{x}$ ではトピックの関連性も捕捉しており，提案手法はトピックの類似性も捉えたと考えられる．

4. おわりに

本研究では，スタイルの類似性を捉える単語ベクトルを教師なし学習することを試みた．具体的には，「発話内の単語のスタイルは一貫している」という仮定に基づいて，ターゲット単語から文脈単語までの位置によって学習する単語ベクトルを調節することで，スタイルの類似性を捉えた単語ベクトルと，統語的/語義的な類似性を捉えた単語ベクトルを同時に学習するモデルを提案した． 実験では，提案モデルが 2 つの性質の異なる類似性をそれぞれ捉えていることを，定量的および定性的に確認した．なお，本研究では定性評価を実施するために，単語ペアに対して人間がスタイルの類似度スコアを付与した新しい評価用データセットを作成した．このスタイルの類似性を捉えられたかを定性的に評価する手法は我々が初めて提案した手法である．

しかし，定性評価の結果，提案したスタイルベクトルはスタイルの類似性に加えてトピックの類似性も捉えていることが明らかになった．今後は，「文書（複数の発話）のトピックは一貫している」という仮定を基に，スタイルとトピックが分離できるようにモデルを拡張する．

謝辞

本研究の一部は，豊田中央研究所および文部科学省科研費 15H01702 の支援を受けた．

参考文献

[1] E. Pavlick and J. Tetreault. An Empirical Analysis of Formality in Online Communication. *Transactions of the Association of Computational Linguistics*, 4:61–74, 2016.

[2] X. Niu and M. Carpuat. Discovering Stylistic Variations in Distributional Vector Space Models via Lexical Paraphrases. In *Proceedings of the Workshop on Stylistic Variation*, pages 20–27, 2017.

[3] D. Wang, N. Jojic, C. Brockett, and E. Nyberg. Steering Output Style and Topic in Neural Response Generation. In *EMNLP2017*, pages 2130–2140, 2017.

[4] E. Pavlick and A. Nenkova. Inducing Lexical Style Properties for Paraphrase and Genre Differentiation. In *NAACL-HLT2015*, pages 218–224, 2015.

[5] L. Flekova, D. PreoŢiuc-Pietro, and L. Ungar. Exploring Stylistic Variation with Age and Income on Twitter. In *ACL2016*, pages 313–319, 2016.

[6] D. Preotiuc-Pietro, W. Xu, and L. H. Ungar. Discovering User Attribute Stylistic Differences via Paraphrasing. In *AAAI2016*, pages 3030–3037, 2016.

[7] R. Sennrich, B. Haddow, and A. Birch. Controlling Politeness in Neural Machine Translation via Side Constraints. In *NAACL-HLT2016*, pages 35–40, 2016.

[8] X. Niu, M. Martindale, and M. Carpuat. A Study of Style in Machine Translation: Controlling the Formality of Machine Translation Output. In *EMNLP2017*, pages 2804–2809, 2017.

[9] W. Xu. From Shakespeare to Twitter: What are Language Styles all about? In *Proceedings of the Workshop on Stylistic Variation*, pages 1–9, 2017.

[10] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient Estimation of Word Representations in Vector Space. In *ICLR2013*, 2013.

[11] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. Distributed Representations of Words and Phrases and Their Compositionality. In *NIPS2013*, pages 3111–3119, 2013.

[12] C. Miyazaki, T. Hirano, R. Higashinaka, T. Makino, and Y. Matsuo. Automatic Conversion of Sentence-end Expressions for Utterance Characterization of Dialogue Systems. In *PACLIC2015*, pages 307–314, 2015.

[13] J. Li, M. Galley, C. Brockett, G. Spithourakis, J. Gao, and B. Dolan. A persona-based neural conversation model. In *Proc. ACL2016*, pages 994–1003, 2016.

[14] D. Gerz, I. Vulic, F. Hill, R. Reichart, and A. Korhonen. SimVerb-3500: A Large-Scale Evaluation Set of Verb Similarity. In *EMNLP2016*, pages 2173–2182, 2016.

[15] S. Kinsui. *Vaacharu nihongo: yakuwari-go no nazo (In Japanese)*. Tokyo, Japan: Iwanami, 2003.

[16] Y. Sakaizawa and M. Komachi. Construction of a Japanese Word Similarity Dataset. In *LREC2018*, 2018.