ARG WI2 No.11, 2017
WI2-2017-26

機能語を次元とするベクトル空間への射影による コーパス間での単語コンテキスト差異の検出指標

木村 和哉　　山西 良典　　西原 陽子

立命館大学情報理工学部

{*sj0035ei@ed, ryama@media, nisihara@fc*}*.ritsumei.ac.jp*

概要　現在，様々な単語の分散表現を獲得する手法がある．fastText を始めとする単語分散表現を取得する多くの手法では文書を入力し，隠れ層と出力層からなる 2 層のニューラルネットワークで学習を行い， 隠れ層における各単語の重みを抽出することで各単語の分散表現を獲得することができる．これらの手法では，学習ごと，つまり学習するコーパスごとで得られた単語ベクトルモデルの各要素に関連性はない．本稿では， 機能語を次元とするベクトル空間へ各学習で得られた単語ベクトルを射影することで，コーパスによって異なる単語のコンテキストを検出する指標を提案する．単語のコンテキストが概ねのコーパスで変化しないと考えられる機能語と各単語との類似度を要素とするベクトルを生成する．生成したベクトルを参照して， 同一単語について異なるモデル間でのベクトル類似度を提案指標とする．提案指標の妥当性を提案指標と主観評価による単語コンテキストの差異の評価の比較によって評価した．

キーワード　単語分散表現，自然言語処理の応用，語義曖昧性の発見

1　はじめに

単語のコンテキストや単語に対する感性は，ドメインによって変化することがある．例えば，歌詞中の “雨” という単語は,「悲しい」というイメージや「涙」の比喩表現などに多く使われる．一方で，ニュース記事中では，“雨” という単語は多くの場合で天候情報を表しており,「悲しい」という印象を表現するために用いられることは少ない．このような，同一単語の語義や感性は単語が用いられるドメインによって変化するが，異なるドメイン間で異なるコンテキストを有する単語を把握することは難しい．

現在，単語分散表現 [1, 2, 3] は，様々な自然言語処理タスクにおいて高い有用性を示している．2016 年に発表された最新の単語分散表現を獲得手法である fastText [4] は，単語中の subword 情報を用いることで単語間の類似性を把握する性質や学習の速さから既に様々なサービス（サイバーエージェントの楽曲聴き放題サービス AWA [5] やリクルートテクノロジーズのレコメンド手法 [1]など）で応用されている．しかし， fastText を始めとする既存の単語分散表現の獲得手法で学習される単語ベクトルモデルは， 他のコーパスを学習して得られた別の単語ベクトルモデルと関連性をもつようには設計されて作られてはいない．そのため，ベクトルの各要素間には関連性が存在せず，複数のベクトルモデル間のベクトルモデル同士をそのまま比較することはできない．ベクトルモデル間をこえて，単語ベクトルの差異を検出することが可能になれば，異なるコーパス，つまりドメインでの同一単語間についてのコンテキストを検知可能になることが期待される．

本稿では，上述の問題を解決するために，概ねのコーパスで単語のコンテキストに変化がないと考えられる機能語に着目する．機能語を基準とし，各ベクトルモデル中の任意の単語と機能語との距離を計算し，この距離を二次的な単語ベクトルとして用いることで，異なるベクトルモデル間での関連性を持たせる．つまり，機能語を次元とするベクトル空間へ各ベクトルモデルの学習で得られた単語ベクトルを射影する．本稿では，このアイデアの妥当性について基礎的な検討を行う．

2　関連研究・関連知識

2.1　語義曖昧性解消

本稿では，単語のコンテキストの差異を検出する指標を提案している．単語コンテキストの判定という観点からは，自然言語処理分野における語義曖昧性解消（Word Sense Disambiguation: WSD）や語義推定（Word Sense Induction: WSI）が関連研究として挙げられる．

WSD とは，複数の語義を持つ単語が文章中に出現した際に，どの語義を表しているかを判断する代表的な自然言語処理タスクの一つであり，語義の分布は品詞と比べてドメインに強く依存することが報告されている [6, 7].

WSD は教師あり学習と教師なし学習の 2 種類に大別される．教師あり学習では人手で用意された教師データを利用して，Support Vector Machine（SVM）などの機械学習手法によって学習を行う [8, 9]．教師あり学習は高い精度で多義語の語義を推定することが可能である

[1]https://www.slideshare.net/recruitcojp/ss-56150629

が，学習データを用意するには多大なコストがかかる．そのため，近年では，教師なし学習による WSD の研究も盛んに取り組まれている．例えば，Pedersen らは WSD の対象語の語義と周辺単語の語義との間の意味的類似性を計算し，適切な語義を選択する手法を提案している [10]．また，佐々木らは多義語の周辺に現れる語義の分布を利用する周辺語義モデルを提案している [11]．

一方，辞書の語義を使用せずに，文脈から多義語をクラスタリングする手法は WSD とは区別され，WSI として議論されることが多い．Agirre らは，多義語の周辺単語の共起情報を基にクラスタリングを行い，語義を判断する手法を報告している [12]．確率的なモデルで WSI を行った研究としては，周辺の文脈から複数種類の素性を抽出し，それらを組み合わせる確率的な手法で Brody らが有効性を示している [13]．

本稿では，WSI に至る一つのアプローチとして単語分散表現を用いて，語義曖昧性の有無の検出を目指す．これは，WSI でターゲットとするべき単語を検出することを目的としており，従来の WSI 研究の前段階のタスクに位置付けられる．

2.2 単語分散表現

近年，多くの自然言語処理分野のタスクにおいて高い性能を示している単語モデル化手法として，Word Vector（単語分散表現）がある．単語分散表現は単語の意味表現を，ニューラルネットワークの学習時に得られる畳み込みベクトルを単語固有のベクトルとして応用したものであり，単語同士の関係性を数理的に演算可能であるという性質をもつ [1, 2]．例えば，“君” に対して，“あなた” のように意味的に関連が強い単語のベクトルは，他の単語に比べてコサイン類似度が高くなる傾向にある．

単語分散表現を獲得するためには，主に CBoW モデルと Skip-gram のモデル [2] が利用される．CBoW (Continuous Bag-of-Words) モデルでは，単語分散表現としてベクトル化したい単語 wt の前後に存在する $2k$ 個の単語を文脈と呼び，この文脈の Bag-of-Words 表現が入力に相当する．単語 wt が出力層に出現する確率を求めるように，ニューラルネットワークの重みを調整しながら学習を進める．一方で，Skip-gram モデルでは，Word Vector 化したい単語 wt を入力層に与え，出力層では文脈中に出現する他の単語 $wt+k$ を推定できるように学習を行う．

単語分散表現を獲得するための手法としては主に word2vec [3] が用いられてきたが，2016 年にはこの手法を発展させた fastText [4] が提案された．word2vec と fastText の大きな違いの一つは，従来モデルとは異なり，単語中の subword 情報を用いることで活用形をまとめられる特性を備えている点である．例えば，“go,” “goes,” “gone, ” “going” などは全て “go” の活用形であるが，従来の word2vec では単語そのものを独立したシンボルとして扱っていたため，これらの単語の関係性は文脈の類似性のみによって構成されていた．fasText では，subword 情報を学習時の単位として用いることで，これらの単語中に共通した “go” という 2-gram が共通して存在することをベクトル学習に反映させる．これにより，共通した n-gram 部分を持つ単語同士は，類似したベクトルをもつように学習される．この特性は，活用形や省略形を捉える問題においては高い性能を示す一方で，共通した n-gram を持つが意味が全く異なる単語同士であっても類似したベクトルをもってしまうという弊害も報告されている．

3 提案手法

提案手法の基本的な考え方は， 図 1 のように，選定した機能語ベクトルと単語ベクトルモデルとの類似度から新たに機能語を基準とした二次的なベクトルモデルを生成する．本稿では，日本語の品詞が「助詞」「助動詞」「接続詞」に該当する単語を機能語として扱った．

3.1 単語ベクトルの獲得

コーパス m の学習で得られたベクトルモデル V_m は，式 (1) で表現される．

$$\{V_m \mid \boldsymbol{v_m}(w_1), \boldsymbol{v_m}(w_2), \ldots, \boldsymbol{v_m}(w_j)\}, \tag{1}$$

ここで，$\boldsymbol{v_m}(w_j)$ は，コーパス m の学習によって得られる単語 w_j のベクトルを示す．

3.2 基準単語の選定

用意したコーパス m と n のいずれにも多数含まれる機能語を基準単語として選択し，ベクトル空間の射影に用いる．用意したコーパスを分かち書きする際に，単語と品詞，単語頻度を記述した単語情報リストを生成する．コーパス m に含まれる単語 w の頻度 $f(w)$ の集合は，式 (2) で示される．

$$\{F_m \mid f_m(w_1), f_m(w_2), \ldots, f_m(w_j)\}, \tag{2}$$

ここで，j は単語のインデックスを示す．

異なるコーパス m と n に含まれる機能語 w_j について，式 (3) に従って各コーパス中での出現頻度の積 $CoFq$ を算出する．

$$CoFq_{m,n}(w_j) = f_m(w_j) \times f_n(w_j). \tag{3}$$

この，$CoFq$ の上位 k 個を基準単語として選出する．これにより得られる基準単語の集合 $S_{m,n}$ は，式 (4) で示される．

$$\{S_{m,n} \mid s_1, s_2, \ldots, s_k\}, \tag{4}$$

ここで，s_k は $CoFq_{m,n}$ の上位 k 番目の機能語を表す．

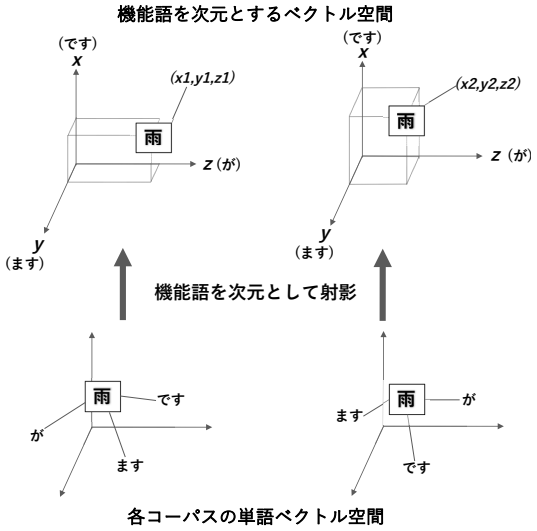

図 1 機能語を基準とする空間への単語ベクトルの射影

3.3 機能語を次元とするベクトル空間への射影

式 (4) で得られたコーパス m と n に関する基準単語についてのベクトル集合 $SV_{m,n}$ を作成する.

$$\{SV_{m,n} \mid \boldsymbol{v_{m,n}(s_1)}, \boldsymbol{v_{m,n}(s_2)}, \ldots, \boldsymbol{v_{m,n}(s_k)}\}, \quad (5)$$

ここで，$\boldsymbol{v_{m,n}(s_k)}$ は k 番目の基準単語 s のベクトルを表す.

このベクトル集合をもとにして，図 1 のように，k 個の基準単語を次元とする新たなベクトルモデルを生成する. 各単語について，基準単語との相対的な関係を要素とする新たなベクトルモデルの各要素にはコーパスを学習して生成されたベクトルモデルの単語と各基準単語のベクトルとのコサイン類似度を用いる．コーパス m で作成された単語 w_j のベクトルを，コーパス m と n を比較する上での基準単語を用いたベクトル空間へと射影する際のベクトルの要素 $rv^m(w_j)$ は，式 (6) で示される. 式 (6) における Sim は 2 つのベクトルのコサイン類似度を表す.

$$rv^m_k(w_j) = Sim(\boldsymbol{v_m(w_j)}, \boldsymbol{v_{m,n}(s_k)}). \quad (6)$$

$SV_{m,n}$ 中の全ての基準単語について式 (6) に従って，$rv^m(w_j)$ を算出する．この $rv^m(w_j)$ を要素とすることで，コーパス m と n についての基準単語を次元とするベクトル空間へ単語 w_j のベクトルを射影する．射影された w_j のベクトル $\boldsymbol{RV^m(w_j)}$ は，式 (6) で得られる $rv^m_k(w_j)$ を用いて，式 (7) で示される.

$$\boldsymbol{RV^m(w_j)} = (rv^m_1(w_j), rv^m_2(w_j), \ldots, rv^m_k(w_j)). \quad (7)$$

3.4 コンテキスト差異を検出するための指標

機能語を次元とするベクトル空間への射影によって生成されたコーパス m と n それぞれのベクトルモデルから $\boldsymbol{RV(w_j)}$ を求める．これを用いて，式 (8) のように，同一単語のベクトルのコサイン類似度をコーパス間での単語 w_j についてのコーパス m と n の間でのコンテキスト差異を検出するための指標 $I_{m,n}(w_j)$ とする.

$$I_{m,n}(w_j) = Sim(\boldsymbol{RV^m(w_j)}, \boldsymbol{RV^n(w_j)}) \quad (8)$$

$I_{m,n}(w_j)$ が高い値を示す時，コーパス m と n においては各コーパス中の基準単語からの w_j に対する相対的な位置が近似していることを示すため，コーパス間でのコンテキストには差異が少ないことが示唆される．一方で，$I_{m,n}(w_j)$ が高い値を示す場合には，コーパス m と n においては各コーパス中の基準単語からの w_j に対する相対的な位置が離れていることを示すため，各コーパスでの w_j のコンテキストに差異が存在することが予測される.

4 評価実験

4.1 実験環境

分かち書きされたコーパスを学習させることで単語ベクトルモデルを得る．分かち書きの際の形態素解析器と辞書には，それぞれ MeCab [14] と NEologd [2]を用いた. 本稿では，単語のコンテキスト差異の検証を行うドメインとして，歌詞とニュース記事を用意した．3,182,766 行（フレーズ）から成る歌詞[3]コーパスと，4,830,623 文から成るニュース記事コーパスの 2 種類のコーパスを用意し，それぞれのコーパスで単語ベクトルモデルを学習した．単語分散表現の獲得は fastText を利用し，モデル学習には Skip-gram を用いた．このとき，単語出現頻度が 5 以下のものは学習させないようにした.

[2]https://github.com/neologd/mecab-unidic-neologd/
[3]http://www.utamap.com/

表 1 実験に使用する単語

高頻度（H）	中頻度（M）	低頻度（L）
人	加速度	おちょこ
心	ライセンス	手塩
風	熊	舞踊
夜	範囲	能登半島
君	ガードレール	利き腕
僕	カプセル	瀬戸内海
明日	死体	男衆
何	博多	天城
胸	抜群	蚊帳
世界	せっかち	さかり
涙	ねずみ	千島
空	撃	鮭
誰	ブラウン管	菜
私	画	寒天
俺	世界地図	消印
夢	九月	青森駅
愛	グラビア	赤城山
今日	フライデー	在所
恋	真顔	真情
あなた	拍手喝采	お札

表 2 出現頻度別での $I_{m,n}(w_j)$ の値が高い上位 3 件

	単語	$I_{m,n}(w_j)$
H	あなた	0.9379
	恋	0.9232
	今日	0.9210
M	拍手喝采	0.9300
	真顔	0.9074
	フライデー	0.9060
L	札	0.9108
	真情	0.9101
	在所	0.9070

評価実験に用いた単語は，歌詞コーパス中での単語出現頻度の上位 20 件（H グループ），中頻度 20 件（M グループ），下位 20 件（L グループ）を用いた．学習時のパラメータを考慮し，実験に用いる L グループについては出現数が 6 以上のものから 20 単語を用いた．表 1 に，本実験で使用した単語 60 件を示す．表 1 に示した 60 単語について，3.4 節で提案した指標 $I_{m,n}(w_j)$ の値とそれぞれのコーパス中で各単語が用いられた文を比較し，考察した．

4.2 実験結果

H，M，L のグループそれぞれ 20 件の単語のうち，提案指標 $I_{m,n}(w_j)$ の値が上位 3 件，下位 3 件となった単語を，表 2 と表 3 にそれぞれ示す．表 3 から，提案指標において値が低かったもののうち，H グループの平均は 0.8507、M グループの平均は 0.7701、L グループの平均は 0.8284 で，M グループが僅かに H グループや L グループに比べて類似度が低い結果となった．これは高頻度で出現する単語に比べて，中頻度の単語の方が学習するデータ数が少なく，ドメインの影響を受けやすいためと考えられる．

表 3 出現頻度別での $I_{m,n}(w_j)$ が低い上位 3 件

	単語	$I_{m,n}(w_j)$
H	人	0.8335
	心	0.8505
	風	0.8683
M	加速度	0.7244
	ライセンス	0.7688
	熊	0.8172
L	おちょこ	0.8050
	手塩	0.8394
	舞踊	0.8410

4.2.1 文脈上の違いに関する考察

表 3 の単語が含まれる文を各コーパスから抽出して考察した．「人」が用いられる文は，ニュース記事中では，“うち海外が 1 万人以上を占める。” や “ニューヨーク連邦地裁は３０日、昨年９月の米中枢同時テロで逮捕、拘束されたヨルダン人学生への米捜査当局の対応は米国憲法違反などとして、起訴を無効とする決定を下した。” のように，名詞の接尾辞として使われる場合が大半を占めた．一方で，歌詞中では “いけない人じゃないのにどうして” や “あの人はもう私のことを” [4]のように，名詞一般として用いられることが多かった．表層系が同じであったとしても，品詞詳細が異なる単語同士について提案指標は低い値を示し，単語の用いられ方が異なることを示唆する結果となった．実験に用いた 60 単語の中で提案指標が最も低い値を示した「加速度」は，デジタル大辞泉 [5]によると，「一定時間内の速度の変化の割合」と「物事の変化の速さがしだいに増していくこと」の 2 種類の意味がある．ニュース記事中では，“ダミー人形の腰にかかる大きな加速度を計測したが、生命には別条ない程度の衝撃だった。” や “再現実験により、発射時の加速度で基板が変形して金具が接触、分離信号が出ることを確認した。” といったように，「一定時間内の速度の変化の割合」を意味する文が大半であった．一方で，歌詞中では “涙が加速度つけて走るわ” [6]や “恋が加速度つけて” [7]のように比喩的に用いられることが多く，語義としては「物事の変化の速さがしだいに増していくこと」を意味する文が多い結果となった．提案指標の低い値は，これらの語義の違いが反映された結果と考えられる．

次に，表 2 の単語が含まれる文を各コーパスから抽出して，考察した．「あなた」ではニュース記事中で “「新聞協会の皆さん、あなた方は『怪しい公人たち』の側につくのですか。” や “チラベルト選手が「あなたたちは

[4]曲名: ウナ・セラ・ディ東京 テレサ・テン，作詞者: 岩谷時子
[5]https://kotobank.jp/dictionary/daijisen/
[6]曲名: 迷宮のアンドローラ，作詞者: 松本隆
[7]曲名: NECESSARY，作詞者: 五十嵐充

表 4 「人」の各ベクトルモデルの類似単語

記事コーパス	歌詞コーパス
１００人	人達
全員	たゞ
ら	お人よし
人数	仮の姿
数人	人出

表 5 「加速度」の各ベクトルモデルでのベクトル類似語

ニュース記事コーパス	歌詞コーパス
重力加速度	最高速度
加速度計	速度
基準地震動	物凄い
地震動	ぐんぐん
上下方向	スピードアップ

まだ若い。” のように代名詞として使われていた．また，歌詞中でも，“希望という名のあなたをたずねて” [8]や “あなたの口から出てくるなんて心うたがうわ” [9]のように同様に代名詞として使われていた．また，表 2 の M や L グループの単語を見ても,「拍手喝采」についてのニュース記事中の “立ち見も出た満場の客席からは拍手喝采が送られた。” と歌詞中の “本来は俺に拍手喝采” [10]のように類似した文脈で使用された事例が多く確認された．コーパス間で語義が類似した単語では提案指標が高い値を示すことが確認された．

本稿のベクトル学習時に用いた fastText の Skip-gram モデルでは，ある単語を入力した時に，その周辺にどのような単語が現れやすいか予測するという性質をもつ．上記の「人」や「加速度」のように，使われ方がニュース記事と歌詞で大きく異なる単語は学習する際の周辺語が大きく異なり，逆に「あなた」や「拍手喝采」といった単語は各コーパスで周辺語が類似していたと考えられる．提案指標では，使われ方がドメインに依存しない機能語との相対的な関係をもとに，これらの周辺語の違いを検出したたため，コンテキスト差異の有無を示唆可能であったと考えられる．

4.2.2 主観評価と文脈における考察

考察対象とした単語については，主観評価でコンテキストの類似度を評価した．各コーパスのベクトルモデルから各単語のベクトルが類似した単語 10 件を 1 セット（ベクトル類似語）として見比べさせ，類似単語のセットが類似した概念であるかを 4 段階評価で評価させた．被験者には，20 代の情報理工学部に所属する大学生 27 名を用意した．

H グループの中で最も提案指標が最も低い値を示した「人」では，各ドメインから得られた 2 種類のベクトル類似語について，被験者の約 9 割が類似していないと回答した．表 4 に,「人」の各ベクトルモデルにおけるベクトル類似語のうち，ベクトル類似度の上位 5 単語を示す．ニュース記事コーパスを学習したベクトルモデルでは「100 人」のように接尾辞と使われる単語が「人」のベクトル類似語されているのに対して，歌詞コーパスを学習したベクトルモデルでは接尾辞ではない使われ方をした単語がベクトル類似語として出力された．

しかしながら，M グループで提案指標が最も低い値を示した「加速度」では，被験者の 6 割以上がベクトル類似度を見て，似ている概念と評価した．表 5 に,「加速度」の各ベクトルモデルにおけるベクトル類似語のうち，ベクトル類似度の上位 5 単語を示す．記事コーパスを学習したベクトルモデルでは「一定時間内の速度の変化の割合」として使用されるであろう単語が数多く見受けられるが，歌詞コーパスを学習したベクトルモデルでは「物事の変化の速さがしだいに増していくこと」の意味として使われる単語が多かった．指標が低い値になったにも関わらず，今回実施した主観評価実験では被験者が辞書的な意味を考慮せずに，一見して「速度に関する単語」と捉えた被験者が多かったためと考えられる．

また，提案指標が最も高い値を示した「あなた」では，8 割以上の被験者が似ている概念であると判断した．表 6 に,「あなた」の各ベクトルモデルにおけるベクトル類似語のうち，ベクトル類似度の上位 5 単語を示す．この単語は，提案指標でも主観評価においてもコンテキストに差異はないと判断されているが，表 6 のように各コーパスで「あなた」のベクトル類似語を見ても,「私」や「君」といった単語が並び，文脈がどちらも近かったこ

表 6 「あなた」の各ベクトルモデルの類似単語

記事コーパス	歌詞コーパス
あなたと	あなたを愛したい
あなたに	あの鐘を鳴らすのはあなた
わたし	君
あなたへ	あなたを想うほど
あなたを忘れない	私

表 7 各コーパスの機能語の類似単語

クエリ	記事コーパス	歌詞コーパス
が	、	は
	の	の
	も	Planet
です	ね	過呼吸
	でし	だ
	けど	面目
に	、	BLACKSUNSHINE
	が	の
	と	連れ出せ

[8]曲名: 希望，作詞者: 藤田敏雄
[9]曲名: 3 年目の浮気，作詞者: 佐々木勉
[10]曲名: Dareder!!!，作詞者: koman’n(Pastel Penguin)

とが伺える．しかし，fastText が subword を考慮している特性上，実験で用いた単語のベクトル類似語には元の単語が含まれているものが多く，被験者は似ているか似ていないの判断が困難であった可能性がある．subword を学習しないベクトルモデルでの評価も検討する必要があると考える．

4.3 機能語を次元とすることの妥当性

本稿では，ドメインによって意味が変化しないと考えられる機能語を基準として，各コーパスを学習して得られたベクトルモデルから機能語で構成される空間に射影している．実験で得られた提案指標の最小値は 0.6462，最大値は 0.9634 となり，想定よりも小さな値域が得られた．この理由として考えられるのは，各コーパスで学習して得られたベクトルモデルにおいて，機能語ベクトル同士の類似度が高いことが挙げられる．表 7 は，各コーパスを学習して得られた単語ベクトルモデルから，機能語のベクトル類似語の類似度上位 3 件をそれぞれ示したものである．“が”，“です”，“に” といった機能語は各ベクトルモデルにおいて，他の機能語との類似度が高いことが確認された．

機能語は，コンテキストに影響しない特徴を持つという点に着目して基準単語として用いたが，他の単語ベクトルとの相対的な距離を測る尺度として考えた時，各機能語が独立性を担保していないため，機能語のみでは基準単語として不十分である可能性が示唆された．表 7 をみると，記号も機能語と同様の性質を持っていることが推察できる．機能語や記号といった単語のベクトル空間上で距離が離れているものを選定するなど，射影空間を構成する基準単語については検討する必要がある．

5 おわりに

本稿では，機能語を次元とするベクトル空間へ各ベクトルモデルの学習で得られた単語ベクトルを射影することにより，異なるコーパス（つまり，ドメイン）での同一単語についてのコンテキストの差異を検知するための指標について基礎的な検討を行った．ニュース記事の歌詞のコーパスを対象として，各単語を機能語 100 件を次元とするベクトル空間へ射影へ行った．単語出現頻度別で選出した単語を提案手法と文中での使われ方とを比較することで，アイデアの妥当性を考察した．

今後は，ドメインによってコンテキストに影響しない特徴を持つ単語が機能語以外に存在するかどうかを調査し基準単語を再検討していくと共に，主観的判断によるコンテキスト差異と相関の高い指標のデザインを目指す．

謝辞

本研究は一部，科学研究費若手研究 B#16K21482 の助成のもと行われた．ニュース記事コーパスについては，Ceek.jp News[11]からの提供を受けた．また，論文中では例示のために歌詞の一部を参照させていただいた．記して謝意を表す．

参考文献

[1] T. Mikolov, W. Yih, and G. Zweig, “Linguistic regularities in continuous space word representations.,” The Proc. of NAACL-HLT, pp.746–751, 2013.

[2] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” The Proc. of Workshop at ICLR, 2013.

[3] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in Advances in Neural Information Processing Systems 26, eds. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, pp.3111–3119, 2013.

[4] P. Bojanowski, E. Grave, A. Joulin, and T. Mikolov, “Enriching word vectors with subword information,” Trans. of the Association for Computational Linguistics, vol.5, pp.135–146, 2017.

[5] 和田計也，福田一郎，“音楽聴き放題サービス awa におけるレコメンド手法の検討 (artist2vec の試み),” 人工知能学会 合同研究会 第 9 回データ指向構成マイニングとシミュレーション研究会，2015.

[6] E. Agirre, and P. Edmonds, Word sense disambiguation: Algorithms and applications, Springer, 2006.

[7] P. Resnik, and D. Yarowsky, “A perspective on word sense disambiguation methods and their evaluation,” Proc. ACL SIGLEX workshop on tagging text with lexical semantics, pp.79–86, 1997.

[8] 新納浩幸，佐々木稔，“k 近傍法とトピックモデルを利用した語義曖昧性解消の領域適応,” 自然言語処理，vol.20，no.5，pp.707–726，2013.

[9] Y.K. Lee, H.T. Ng, and T.K. Chia, “Supervised word sense disambiguation with support vector machines and multiple knowledge sources,” Senseval-3: Third International Workshop on the Evaluation of Systems for the Semantic Analysis of Text, pp.137–140, 2004.

[10] T. Pedersen, S. Banerjee, and S. Patwardhan, “Maximizing semantic relatedness to perform word sense disambiguation,” Research Report UMSI, 2005.

[11] 佐々木悠人，古宮嘉那子，森田一，小谷善行，“周辺語義モデルによる日本語の教師無し語義曖昧性解消,” 情報処理学会研究報告自然言語処理（NL），vol.2014-NL-218，no.3，pp.1–14，2014.

[12] E. Agirre, D. Martinez, O.L. de Lacalle, and A. Soroa, “Two graph-based algrithms for state-of-the-art wsd,” pp.585–593, 2006.

[13] S. Brody, and M. Lapata, “Bayesian word sense induction,” pp.103–111, 2009.

[14] T. Kudo, K. Yamamoto, and Y. Matsumoto, “Applying conditional random fields to japanese morphological analysis,” The Proc. of EMNLP-2004, pp.230–237, 2004.

[11] http://news.ceek.jp/