

画像検索を用いた英単語学習支援

大谷 紀子研究室
0432173 福留 拓也

1. 研究の背景と目的

英単語の学習法のひとつに英単語イメージ学習法がある。英単語イメージとは、ネイティブスピーカーが単語を捉えている感覚である。英単語イメージによる学習は、訳語を暗記する学習法と比較して、豊かな英語表現が身に付く効果的な学習であるとされている。しかし、基本語の意味の理解を深めることが目的であるため、取り扱われる単語は動詞、前置詞、形容詞が 100 語強と少ない[1]。基本語に加えて、より多くの英単語イメージを得るためには、単語の意味を表現する画像が必要であると考えられる。ウェブ上の画像検索では、多種多様な画像の中から検索キーワードに関連する画像を取得することができる。しかし、英単語をキーワードとする画像検索を学習に用いる場合、一般的な利用目的とは異なるため、検索キーワードの品詞によって精度が異なる点と、結果の表示が非効率的である点が問題となる。一般に画像検索エンジンは、画像の周囲に存在する名詞キーワードによって、画像の特徴を検出する[2]。したがって、名詞以外の英単語について学習する場合、検索結果の画像から意味を読み取ることが難しい。また、検索結果の画像は検索エンジン独自のランクに従って表示されるが、英単語イメージ学習では、take の場合「...を連れていく」と「写真を撮る」といった用例別に画像を表示することが望ましい。

本研究では、英単語イメージ学習の支援を目的として、画像検索を用いた単語イメージ取得方法を提案する。英単語イメージは代表的な用例別に取得する。名詞と動詞を研究の対象とし、動詞については、検索キーワードを動名詞化して得られる画像についても考察する。

2. 提案手法

ウェブ上の画像は、画像を含む文書により内容が判別される。画像 I を含む文書が、英単語 T を含む文書と類似しているとき、画像 I は英単語 T の単語イメージを表すと考えられる。英単語 T を含む文書はウェブ検索によって取得し、各文書の内容をベクトルで表現してクラスタリングする[3]。文書をベクトルで表現することで、文書の特徴を数値化し、文書同士の類似度を求めることができる。各クラスタは、英単語 T の代表的な用例ごとの文書集合であると考えられる。画像 J を含む文書のベクトルが、英単語 U のひとつの用例 A を表すクラスタの平均ベクトルと類似するとき、画像 J は英単語 U の用例 A を表現する画像といえる。英単語の代表的な用例ごとに、クラスタと最も類似する文書に含まれる画像を、単語イメージとして抽出する。クラスタと文書の類似度は、クラスタの平均ベクトルと文書ベクトルの類似度に等しい。

3. 実験

3.1. 実験の概要

基本語と、高校英語で学習する専門語[4]について画像を取得し、英単語のイメージとの適合度により提案手法を評価する。基本語として、way, heart, head などの名詞 16 個、take, give, get などの動詞 56 個と、専門語として、bias, collaboration, confession などの名詞 44 語、arrest, avail, cherish などの動詞 18 語を実験に用いる。

3.2 . 実験結果

全体の 49%の英単語について単語イメージの取得に成功した。明確に意味を表現する画像を適合画像、意味の理解を深めることに有効な画像を半適合画像と呼ぶ。表 1 に英単語別の適合画像と半適合画像の取得結果を示す。括弧内の数値は、動詞を動名詞化して単語イメージを取得した結果である。face, eye, mouth など画像として表現しやすい名詞は、明確に意味を表す画像を取得できたが、end, sense, interest など抽象的な意味を表現する名詞は効果的な画像を得られなかった。動詞の場合は、基本語で 36%、専門語で 28%の取得率であったが、動名詞をキーワードとして画像を取得すると、取得率は約 1.5 倍向上した。図 1, 2 に catch と catching の取得画像を示す。図 1 では、catch の目的語となるシマウマが主体になっているが、図 2 では catch が表す動作が主体になっている。類義語動詞の取得例として、図 3 に lifting, 図 4 に raising の取得画像を示す。図 3, 4 を比較すると、類義語動詞のニュアンスの違いを得ることができる。また、英単語の用例別に画像を取得した結果、多くの英単語については、2~3 個の用例別クラスタに対して同一の画像を類似度の高い画像として取得した。

表 1 単語イメージ取得率

		適合	半適合	適合+半適合	単語数	取得率
基本語	名詞	4	7	11	16	69%
	動詞 (ing)	5 (12)	15 (21)	20 (33)	56 (56)	36% (59%)
専門語	名詞	0	25	25	44	57%
	動詞 (ing)	1 (0)	4 (7)	5 (7)	18 (18)	28% (39%)



図 1 catch



図 2 catching



図 3 lifting



図 4 raising

4 . 考察

適合画像が得られた英単語は、全体で 11%にとどまったが、半適合画像を加えると約半数の英単語について単語イメージを取得することができた。英単語イメージ学習の本質は、単語の意味の理解を深めることであり、画像検索を用いた英単語イメージ学習と他の学習法と組み合わせることで、意味の理解を高める効果が期待できる。用例別の画像抽出については、ウェブ検索の対象を絞ることで精度が向上すると予想される。また、学習者のレベルに合わせた英文書を用いて単語イメージを取得することで、より効果的な学習が可能になると考えられる。ウェブ上の情報は多言語に及ぶため、単語イメージ学習を英単語以外の言語への応用することも期待できる。

参考文献

- [1] 田中茂範, 佐藤芳明, 河原清志: イメージでわかる単語帳, NHK 出版 (2007)
- [2] 柳井啓司: 一般画像自動分類の実現へ向けた World Wide Web からの画像知識の獲得, 人工知能学会論文誌 Vol.19 pp. 429-439 (2004)
- [3] 大谷紀子: 情報検索におけるベクトル空間モデルの応用, 武蔵工業大学環境情報学部紀要第五号, pp. 99-109 (2004)
- [4] 水島孝司 監修: 英単語 2400 受験英語からの TOEIC Test, Z 会出版, pp. 458-477 (2004)