国立国語研究所学術情報リポジトリ

絵文字の統語的分類に関する調査

メタデータ	言語: Japanese 出版者: 公開日: 2022-01-07 キーワード (Ja): キーワード (En): Yahoo! Crowdsourcing 作成者: 岡本, 沙紀, 落合, 哉人, OKAMOTO, Saki, OCHIAI, Kanato メールアドレス: 所属:
URL	https://doi.org/10.15084/00003481

絵文字の統語的分類に関する調査

岡本沙紀（東京大学）†
落合哉人（筑波大学）

A survey on Syntactic Classification of Emoji

Saki Okamoto (University of Tokyo)
Kanato Ochiai (University of Tsukuba)

要旨

絵文字は，1999 年ドコモ「i モード」サービス開始以降日本で普及し，G メールや iPhone が日本参入し Unicode に登録されたことから世界中に広まった。以来，絵文字は単なる感情や挿絵を付加する記号の枠を超え，語用論的な研究の対象になっている。しかしながら，これまでの絵文字の用法に関する調査は，小規模なものや，自然言語処理的な手法のものが多かった。本研究では絵文字に馴染み深いと思われる 15~40 歳を対象に，1680 個の絵文字の用法について，「体・用・相のうちどれを表すのに使うことができるか」を複数選択可のアンケートで調査した。その結果，顔の絵文字は体言性が弱く相言性が突出して強いこと，交通手段を用いる絵文字では用言性が顕著に高いこと，また食べ物と動物は共通して強い体言性と弱い相言性が見られたが食べ物の方が用言性が強いことなど，絵文字の意味によって特徴ある分布が見られた。

１．はじめに

絵文字の用途の分類法は、Herring and Dainas's (2017)によって提案されている。「感情の付加」「文体を柔らかくする」「前の文章への反応」「動作」「直前の文章の内容説明」「（実際の）身体動作」「飾り付け」の７つのうち、どの絵文字も担いうる役割として「感情の付加」「直前の文章の内容説明」が挙げられている。前者に関しては主に自然言語処理を用いた研究がよくなされているが、後者について細かく分類した研究はない。絵文字のはたらきが文章中の内容の繰り返しや文章への意味の付加である場合、絵文字に対応する文中の要素が存在し、その統語的な特徴によって分類することができると考えた。

２．手法

Yahoo!クラウドソーシングを利用し，絵文字を伴うテキストベースコミュニケーションに若い頃から親しんでいるであろう 15~40 才を対象に，ある絵文字についてそれが以下のどの項目に当てはまるものか，全て選択してもらった：

表示されない
普段よく使う
事物（ヒトやモノ）を表すのに使うことができる

† okamotos@g.ecc.u-tokyo.ac.jp

様子（ありさま）を表すのに使うことができる
出来事（―すること）を表すのに使うことができる
ポジティブな気分を表すのに使うことができる
ネガティブな気分を表すのに使うことができる
文末に使う

回答はスマートフォンからのアクセスのみを受け付け、アンケートの開設時間を夕方に設定することでより若い人からの回答の割合を高めるよう工夫した。約 1000 個の絵文字について、１つの絵文字あたり２０回答を得た。

３．結果

まず，絵文字そのものを二次元にプロットした。ここで，nominal という軸は，「表示されない」と答えなかった人のうち「事物（ヒトやモノ）を表すのに使うことができる」にチェックを入れた人の割合，verbal は「様子（ありさま）を表すのに使うことができる」に，adjectival は「出来事（―すること）を表すのに使うことができる」にチェックを入れた人の割合を表した。マクロな傾向として人の顔を表す絵文字や食べ物を表す絵文字などが偏っていること，ミクロな傾向としてものを表す絵文字ではビール瓶やマイクなどの身近なアクティビティを連想させる絵文字が verbal さが高く，ベーグルやカップケーキやアボカドやココナツなどのあまり身近でないものは nominal さが高いということがわかった。

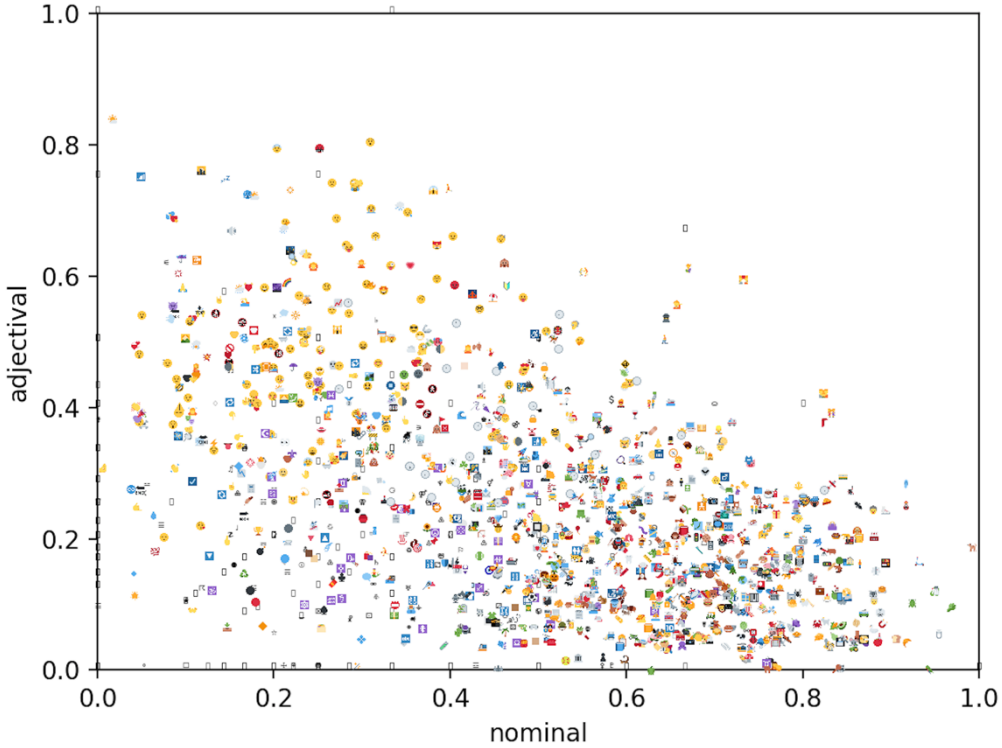

図１　縦軸に用言らしさ、横軸に体言らしさをとった絵文字マッピング

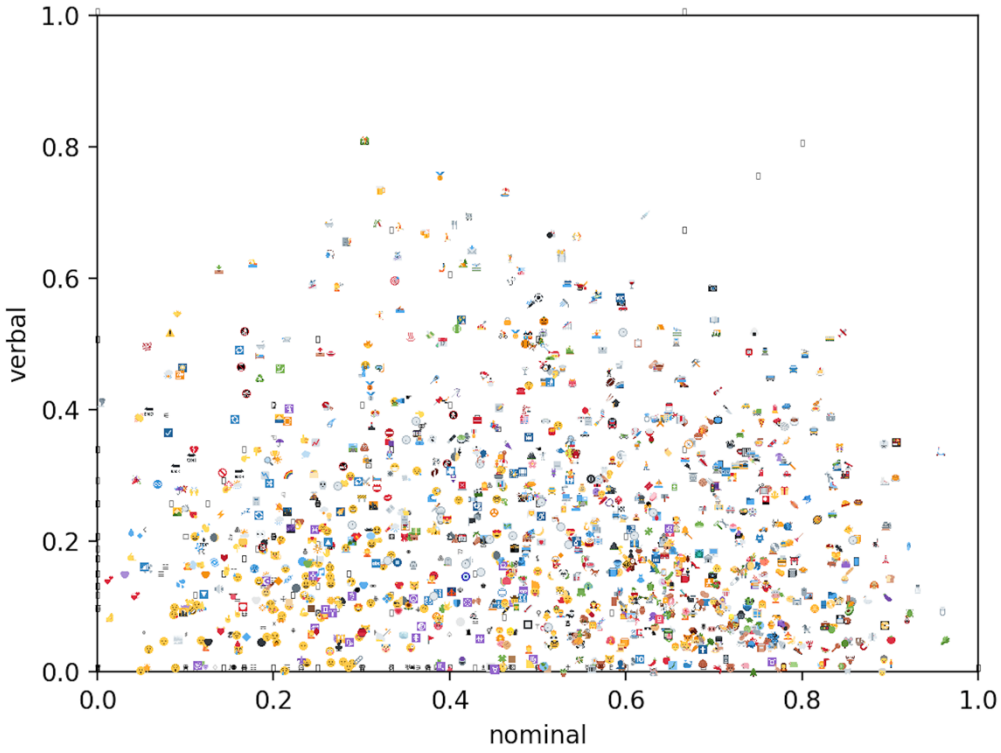

図２　縦軸に相言らしさ、横軸に体言らしさをとった絵文字マッピング

マクロな傾向についてより詳しく見るため、絵文字の主な意味内容ごとに色分けしマッピングを行なった。人の顔や動作（黄色）・動物（赤）・飲食物（緑）・乗り物（青）のそれぞれのグループについて、９０％信頼区間を表す楕円体を同じ色で示した。
人の顔や動作を表す絵文字は明らかに異なる分布を示すこと、その他の３グループは重なる部分も多いもののそれぞれに特徴的な分布を示すことがわかった。

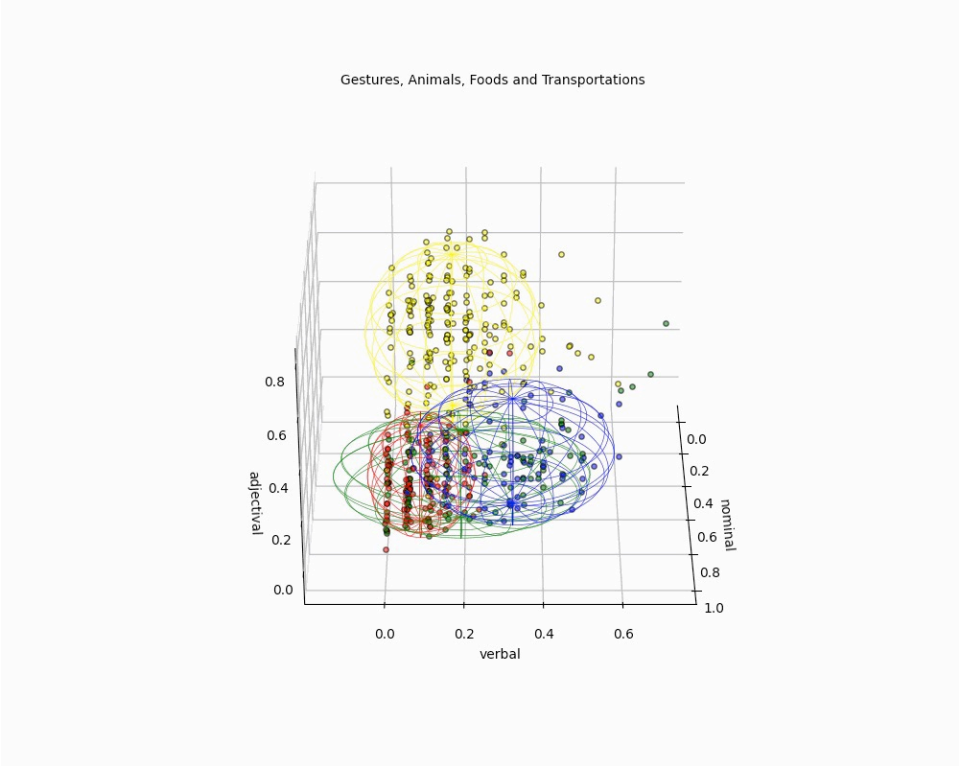

図３ 意味ごとの体用相の割合の分布

４．結論

絵文字を意味ごとに分類することによって，利用者の感じる体言らしさ・用言らしさ・相言らしさのスコアの特徴を分類することができた。身近なアクティビティを連想させる絵文字では用言らしさが高く，ベーグルやカップケーキやアボカドやココナツなどのあまり身近でないものは体言らしさ高いということがわかった。

顔の絵文字は体言性が弱く相言性が突出して強いこと，交通手段を用いる絵文字では用言性が顕著に高いこと，また食べ物と動物は共通して強い体言性と弱い相言性が見られたが食べ物の方が用言性が強いことなど，絵文字の意味によって特徴ある分布が見られた。

今後，絵文字の体用相スコアを用いて絵文字の日本語への適切な置き換えや，「🎤はある歌や声を意味するかカラオケに行くことを意味するか」のような絵文字の意味範囲と使い分けの研究を進めたい。

謝　辞

研究の組み立て方などを一から指導してくださった国立国語研究所の山崎誠先生と浅原正幸先生、Yahoo!クラウドソーシングの実施にあたりましてアドバイスをいただきました国立国語研究所の西内沙恵さん、Python を教えてくれた東京大学大学院数理科学部数学科の青木孔さんに感謝申し上げます。

文　献

Herring, Susan and Dainas, Ashley (2017) "Nice Picture Comment!" Graphicons in Facebook Comment Threads, *Proceedings of the 50th Hawaii International Conference on System Sciences*, pp.2185-2194. http://hdl.handle.net/10125/41419

関連 URL

Yahoo!クラウドソーシング（https://crowdsourcing.yahoo.co.jp/special/owner）