

日本語と英語のツイートの末尾にある不読符号列の比較

中嶋 邦裕^{1,a)} 渡辺 靖彦^{1,b)} 松本 浩輝^{1,c)} 西村 涼^{1,d)} 岡田至弘^{1,e)}

概要: 本研究では、マイクロブログの Twitter に投稿された日本語と英語のテキストメッセージ(ツイート)を比較し、投稿の末尾にある不読符号列がそれぞれどのように用いられているかを調べる。そして、日本語の場合、不特定多数のユーザへのツイートに比べ、特定のユーザへのツイートでは末尾で不読符号列が用いられるケースが増加するのに対し、英語の場合、ほとんど増加しないことを示す。

キーワード: 不読符号列, Twitter, 日英比較, マイクロブログ

A comparative study of unsounded code strings at the end of Japanese and English tweets

NAKAJIMA KUNIHIRO^{1,a)} WATANABE YASUHIKO^{1,b)} MATSUMOTO HIROKI^{1,c)} NISHIMURA RYO^{1,d)}
OKADA YOSHIHIRO^{1,e)}

Abstract: In this study, we compare Japanese and English tweets submitted to Twitter and discuss how we use unsounded code strings at the end of online messages. We first define unsounded codes and unsounded code strings used in Japanese and English text. Next, we compare and discuss the usage of unsounded code strings at the end of tweets to general public and particular persons. Finally, we show that Japanese speakers use unsounded code strings at the end of online messages more frequently to particular persons than to general public while English speakers do not.

Keywords: unsounded code string, Twitter, comparative study of Japanese and English, micro blog

1. はじめに

Twitter は「ツイート」と呼ばれる他者の反応を前提としない発言を投稿するメディアで、コミュニケーションメディアにおける近年の最大の話題の1つである。他者の反応を前提としないとは、相手に割り込みをかけないということである。他者に割り込みをかける場合、用件もなしに割り込みをかけることはのぞましくない。このため、旧来のメディアでは、用件がなければコミュニケーションを開始することがむずかしかった。一方、Twitter では、特に用件がなくても投稿することができ、それをきっかけにコ

ミュニケーションが始まるのが期待できる。このような新しいメディアをわれわれがどのように利用しているのかを明らかにすることは重要である。例えば、Twitter に投稿されたテキストメッセージ(ツイート)について、どんな目的で投稿されたのか、なぜその表現が用いられたのか、などが明らかにすることは重要である。

(例文 1) 課題終わったあああああああああああ
ああ!!!!!!!!!!!!!!

(例文 2) こりゃあ数理統計の課題やる時間ないぞ・・・

(例文 1) と (例文 2) は Twitter に投稿されたツイートの例で、それらの末尾には読み上げるときに発音されない記号や符号が連続して用いられている。この読み上げるときに発音されない記号や符号のことを本研究では不読符号とよぶ。そして、不読符号が 3 個以上連続して用いられている場合、その符号列を不読符号列とよぶ。不読符号列は、

¹ 龍谷大学 理工学部 情報メディア学科

a) t13m071@mail.ryukoku.ac.jp

b) watanabe@rins.ryukoku.ac.jp

c) t100440@mail.ryukoku.ac.jp

d) r_nishimura@afc.ryukoku.ac.jp

e) okada@rins.ryukoku.ac.jp

メールやチャット、マイクログログや Q&A サイトなどのオンラインのコミュニケーションではさかんに用いられている。例えば Yahoo!知恵袋では、25%の回答の末尾で不読符号列が用いられている。これほどさかんに用いられている表現であるのに、この不読符号列についての研究はまだほとんど行われていない。そこでわれわれは、Yahoo!知恵袋に投稿された回答の末尾にある不読符号列の調査を行い、短い回答の末尾で長い不読符号列がさかんに用いられていることを明らかにした [1]。これは、Yahoo!知恵袋には 25 文字未満の回答は投稿できないという投稿文字数制限があり、その制限を回避するために長い不読符号列が用いられていると考えられる。次に、Yahoo!知恵袋の回答と日本語のツイートを対象に、それぞれの末尾にある不読符号列を比較した [2]。その結果、末尾の不読符号列以外の長さが投稿文字数制限 (25 文字) 以上の Yahoo!知恵袋の回答と日本語のツイート、それぞれの末尾にある不読符号列の長さの分布が類似していることを明らかにした。本研究では、英語と日本語のツイートを対象に、それぞれの末尾で用いられている不読符号列を比較する。特に、不読符号列を用いるかどうかに影響を与える要因について検討する。

2. 不読符号列と関連研究

テキストには、読み上げる時には発音されない記号や符号がある。また、読み上げる時には本来発音される文字であっても、用いられている位置や状況によって、発音されないことがある。このような文字や記号や符号をわれわれは不読符号とよぶことにする。本研究では、Twitter に投稿されたツイートの末尾で連続して用いられる場合に発音されないことが多い記号や符号として、英語の場合、

- 英数字以外の ASCII 印字可能文字 (!#\$%&.,:;<=>?@(){} など)

一方、日本語の場合、

- 英数字以外の ASCII 印字可能文字
- 記号 (、 。 , . : ; など)
- ギリシャ文字
- キリル文字
- 罫線

を不読符号とする。そしてそれらが 3 個以上連続してツイートの末尾で用いられている不読符号列について分析を行う。したがって、以下の例の下線部は、本研究では取り扱わない。

(例文 4) おはようー まさかの 6 時おき (笑)

(例文 5) 番長「あいつとおるだけで幸せを感じる。」

(例文 6) Did anyone else wake up as sore as I did??

本研究では、日本語と英語のツイートを比較する。異なる言語の話者の表現や行動については、さまざまな視点からさかんに比較研究が行われている。これらの研究は、以下の 2 つに分類できる。

- 1 つの言語の外国語学習者と母語話者の表現や行動などを比較する
- 異なる言語の母語話者の表現や行動などを直接比較する

本研究は、日本語と英語のツイートの末尾にある不読符号列を直接比較するので、後者の研究に分類される。

外国語学習者と母語話者の表現や行動などを比較する研究として、外国語学習者の中間言語による発話行為について、誉め言葉 [3]、謝罪 [4]、感謝 [5]、丁寧さ [6]、断り [7] などが研究されている。また、外国語学習者が発話でどのようにポーズをおくのかについて、Deschamps はフランス人の英語学習者が英語を話す場合について [8]、石崎は英語・フランス語・中国語・韓国語を母語とする日本語学習者が日本語を話す場合について [9]、それぞれ研究を行っている。堀口は日本語学習者のあいづちについて調べ、学習が進めば、その頻度、種類、適切さが日本語母語話者に近くなると報告している [10]。LoCastro は、日本人の英語学習者が英語を話す場合、あいづちを適切にうてなくて違和感を感じていると報告している [11]。Tera らは、日本語学習者と日本語母語話者を対象に日本語テキスト読解時の視線移動について比較を行っている [12]。音声情報処理の分野では、外国語学習者の音韻や韻律を母語話者と比較して評価する方法が研究されている [13] [14] [15] [16] [17] [18] [19]。

一方、異なる言語の母語話者の表現や行動などを直接比較する研究も行われている。特にあいづちは異なる言語の母語話者間でも比較・分析しやすいため、その頻度の分析を中心にさかんに研究されている。Maynard と Miller は日本語と英語の会話を調査し、英語より日本語であいづちが頻繁に用いられていることを示した [20] [21]。あいづちの頻度に影響を与える要因についての研究も行われている。White は、アメリカ人同士の英語会話に比べて、日本人とアメリカ人の英語会話ではアメリカ人のあいづちが増えることを示した [22]。これは、会話の相手が日本人であるかアメリカ人であるかが、アメリカ人のあいづちの頻度に影響を与える要因の 1 つであることを示している。しかし、あいづちの頻度に影響を与える要因についての研究はまだ少なく、今後研究するべきテーマであると陳は指摘している [23]。そこで本研究でも、不読符号列を用いるかどうかに影響を与える要因について日本語と英語のツイートで比較する。

3. 日本語と英語のツイートの末尾にある不読符号列の調査

本研究では、日本語と英語のツイートの末尾で用いられている不読符号列を比較する。Twitter は、Web ページまたは専用ソフトウェアからツイートと呼ばれる他者の反応を前提としない投稿をするメディアである。本研究では、

表 1 日本語のノーマルツイート、リプライ、リツイートの数 (2012年 11月～12月の3週間)

タイプ	件数および割合	
ノーマル	3,813,164	(53.82%)
リプライ	2,528,642	(35.69%)
リツイート	743,461	(10.49%)
合計	7,085,267	(100.00%)

表 2 末尾に不読符号列がある日本語のノーマルツイート、リプライ、リツイートの数 (2012年 11月～12月の3週間)

タイプ	件数および割合	
ノーマル	356,727	(36.92%)
リプライ	430,294	(44.54%)
リツイート	179,166	(18.54%)
合計	966,187	(100.00%)

この Twitter に 2012 年 11 月から 12 月までの 3 週間に投稿されたツイートから streaming API *1 を用いて取り出した以下のツイートを対象に調査を行う。

- 言語設定で「日本語」を選択したユーザが投稿したツイート (以後、日本語のツイート) 7,085,267 件
- 言語設定で「英語」を選択したユーザが投稿したツイート (以後、英語のツイート) 31,253,241 件

なお、streaming API で取り出すことができるツイートは、Twitter に投稿されたすべてのツイートの 1%以下である。ツイートは以下の 3 種類に分類できる。

リプライ 特定のユーザへの返信。ツイート内に返信相手のユーザ ID(@username) が示されている。

リツイート 他のユーザが投稿したツイートを再投稿したもの。

ノーマルツイート リプライでもリツイートでもないツイート。本研究では、これらのツイートをノーマルツイートとよぶ。Twitter は他者の反応を前提としない投稿をするメディアであるので、ノーマルツイートの多くは、特定のユーザにではなく、不特定多数のユーザに宛てて投稿されたツイートであると考えられる。

表 1 に取り出した日本語のツイート 7,085,267 件の内訳を示す。この 7,085,267 件の日本語のツイートのうち 966,187 件 (13.64%) のツイートの末尾で不読符号列が用いられていた。表 2 に末尾に不読符号列がある日本語のツイート 966,187 件の内訳を示す。一方、表 3 に取り出した英語のツイート 31,253,241 件の内訳を示す。この 31,253,241 件の英語のツイートのうち 3,270,821 件 (10.47%) のツイートの末尾で不読符号列が用いられていた。表 4 に末尾に不読符号列がある英語のツイート 3,270,821 件の内訳を示す。

表 4 に示すように、英語のリツイートでは、末尾に不読符号列があるものが多い。特に、長さが 140 文字であるリ

表 3 英語のノーマルツイート、リプライ、リツイートの数 (2012年 11月～12月の3週間)

タイプ	件数および割合	
ノーマル	16,023,795	(51.27%)
リプライ	8,267,646	(26.45%)
リツイート	6,961,800	(22.28%)
合計	31,253,241	(100.00%)

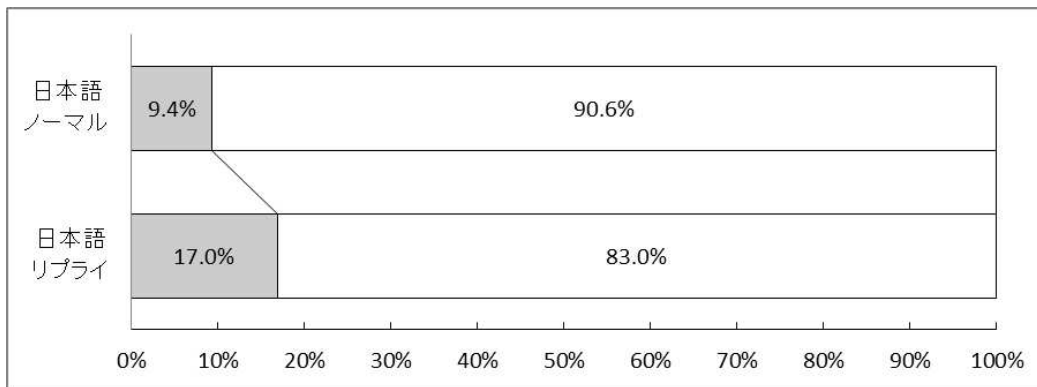
表 4 末尾に不読符号列がある英語のノーマルツイート、リプライ、リツイートの数 (2012年 11月～12月の3週間)

タイプ	件数および割合	
ノーマル	1,121,952	(34.30%)
リプライ	632,298	(19.33%)
リツイート	1,516,571	(46.37%)
合計	3,270,821	(100.00%)

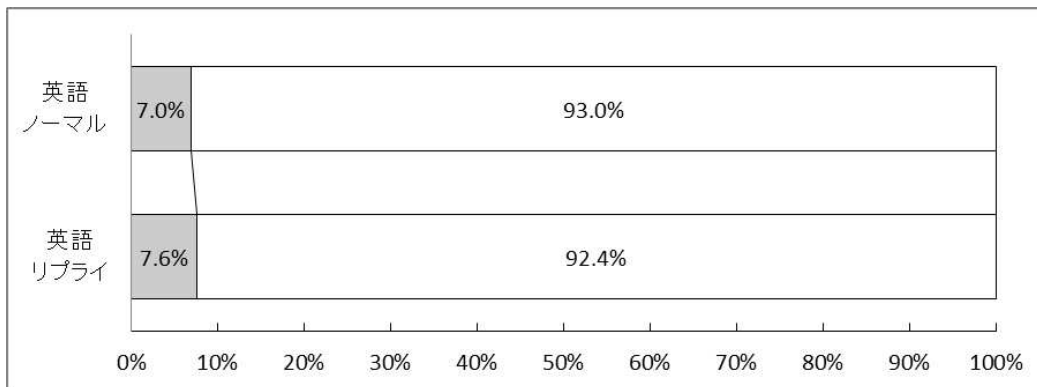
ツイートは 1,249,648 件あった。これは、リツイート情報を追加したため文字数制限 (140 文字) を超過したテキストをリツイート API が切り詰め、その末尾に「…」を自動的に追加するためと考えられる。そもそも、リツイートは他のユーザが投稿したツイートが再投稿されたものである。したがって、リツイートの末尾にある不読符号列は、その投稿者が書いたものではなく、他のユーザやリツイート API が書いたものである。そこで、リツイートは本研究の調査対象から除くことにする。

Twitter は、他者の反応を前提としない投稿をするメディアである。このため、ノーマルツイートの多くは、特定のユーザにではなく、不特定多数のユーザに宛てて投稿されたツイートであると考えられる。一方、リプライは特定の相手に宛てて投稿されたツイートである。そこでまず、ノーマルツイートとリプライでは末尾の不読符号列の用いられ方に違いがあるかを調べる。図 1 に日本語と英語のノーマルツイートとリプライで末尾に不読符号列があるものの割合を示す。図 1 に示すように、日本語のツイートの場合、末尾に不読符号列がある割合はノーマルツイートでは 9.4%、リプライでは 17.0%と大きく異なる。一方、英語のツイートの場合、末尾に不読符号列がある割合はノーマルツイートでは 7.0%、リプライでは 7.6%とほとんど同じである。ノーマルツイートに比べ、日本語のリプライの末尾で不読符号列がさかんに用いられているのは、リプライが特定のユーザに宛てたツイートであるためと考えられる。不読符号列ではなく顔文字についてであるが、日本語の電子メールでは間柄が親しいほどより多くの顔文字が用いられる傾向があると加藤らは報告している [24]。これらのことから、言語設定で日本語を選択したユーザの場合、ツイートが特定のユーザに宛てて投稿するものかどうかは、ツイートの末尾で不読符号列を用いるかどうかに影響があると考えられる。一方、言語設定で英語を選択

*1 <https://dev.twitter.com/docs/streaming-apis>



(a) 末尾に不読符号列がある日本語のノーマルツイートとリプライの割合



(b) 末尾に不読符号列がある英語のノーマルツイートとリプライの割合

図 1 末尾に不読符号列がある日本語と英語のノーマルツイートとリプライの割合

したユーザの場合、ツイートが特定のユーザに宛てて投稿するものであるかどうかは、ツイートの末尾で不読符号列を用いるかどうかにより影響がないと考えられる。したがって、日本語を選択したユーザにとっては、特定のユーザ宛てのツイートであるかどうかは、その末尾で不読符号列を用いるかどうかに影響を与える要因であるが、英語を選択したユーザにとってはそうではないと考えられる。

次に、末尾の不読符号列の有無がツイートの長さを与える影響について検討する。図2と図4にTwitterに投稿されたツイート(リツイートを除く)について以下の長さの累積相対度数分布を示す。

- すべてのツイート(リツイートを除く)の長さ
- 不読符号列が末尾にあるツイート(リツイートを除く)の長さ
- 不読符号列が末尾にあるツイート(リツイートを除く)の末尾の不読符号列の長さ
- 不読符号列が末尾にあるツイート(リツイートを除く)の末尾の不読符号列以外の長さ

なお本研究では、マルチバイト文字とシングルバイト文字を区別せず、それぞれの文字1個を1文字としてツイートおよび不読符号列の長さを測定した。図2に示すように、日本語のツイートの末尾で用いられる不読符号列のうち91%が8文字以下で、さらに58%が5文字以下である。一

方、図4に示すように、英語のツイートの末尾で用いられる不読符号列のうち90%が5文字以下である。したがって、日本語のツイートでも英語のツイートでも、その末尾で用いられる不読符号列は5文字以下のものが多い。長さが6文字以上の不読符号列は、日本語のツイートの末尾ではしばしば用いられるが、英語のツイートの末尾ではあまり用いられない。

図3と図5は不読符号列が末尾にあるツイート(リツイートを除く)について、

- ツイートの末尾の不読符号列の長さ
- 末尾の不読符号列以外の長さ

の関連性を示すヒートマップである。図3と図5のヒートマップで濃い色で示されているのは、ツイートの末尾の不読符号列とそれ以外の長さの組み合わせで出現頻度の高いものである。図3と図5に示すように、ツイートの末尾の不読符号列とそれ以外の長さの組み合わせはさまざまである。さらに図2と図4から、日本語のツイートでも英語のツイートでも、末尾に不読符号列があるツイートの長さの分布は、Twitterに投稿されたすべてのツイートの長さの分布に近いことがわかる。これらのことから、日本語のツイートでも英語のツイートでも、末尾の不読符号列の有無がツイートの長さを与える影響は小さいと考えられる。さらに詳しく検討するため、末尾の不読符号列の有無がノー

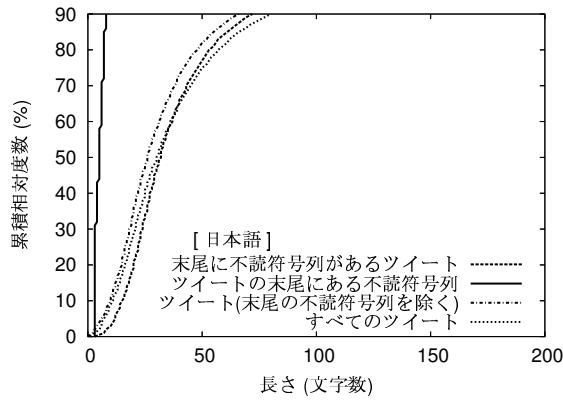


図 2 Twitter に投稿されたすべての日本語のツイート、不読符号列が末尾にある日本語のツイート、そして日本語のツイートの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

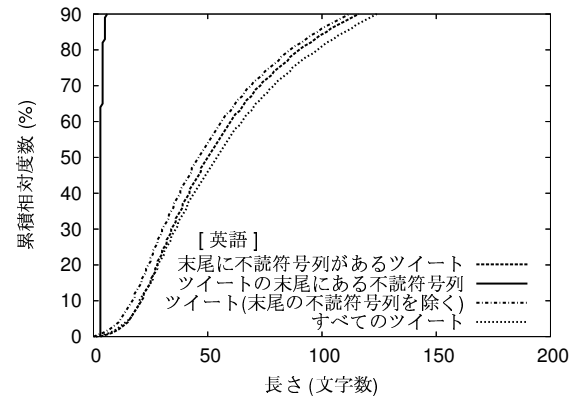


図 4 Twitter に投稿されたすべての英語のツイート、不読符号列が末尾にある英語のツイート、そして英語のツイートの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

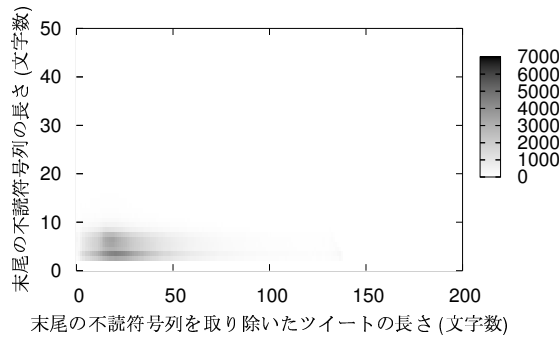


図 3 Twitter に投稿された不読符号列が末尾にある日本語のツイートについて、ツイートの末尾にある不読符号列の長さとその不読符号列を除いたツイートの長さの関連性を示すヒートマップ

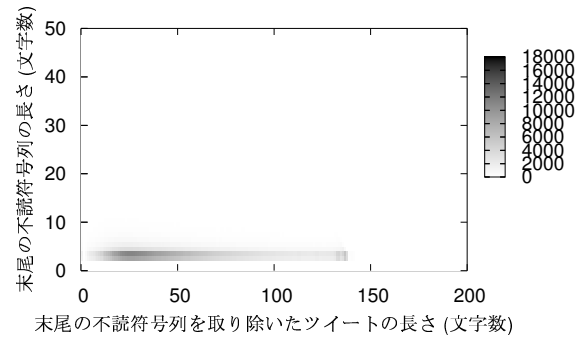


図 5 Twitter に投稿された不読符号列が末尾にある英語のツイートについて、ツイートの末尾にある不読符号列の長さとその不読符号列を除いたツイートの長さの関連性を示すヒートマップ

マルチツイートとリプライの長さを与える影響について調べる。

図 6 と図 8 は日本語と英語のノーマルツイートについて以下の長さの累積相対度数分布を示す。

- すべてのノーマルツイートの長さ
- 不読符号列が末尾にあるノーマルツイートの長さ
- 不読符号列が末尾にあるノーマルツイートの末尾の不読符号列の長さ
- 不読符号列が末尾にあるノーマルツイートの末尾の不読符号列以外の長さ

一方、図 7 と図 9 はリプライについて以下の長さの累積相対度数分布を示す。

- すべてのリプライの長さ
- 不読符号列が末尾にあるリプライの長さ
- 不読符号列が末尾にあるリプライの末尾の不読符号列の長さ
- 不読符号列が末尾にあるリプライの末尾の不読符号列以外の長さ

図 6 と図 7 に示すように、日本語のノーマルツイートとリプライの末尾の不読符号列の長さの分布は類似している。さらに図 6 から、末尾に不読符号列がある日本語のノーマ

ルツイートの長さの分布は、Twitter に投稿されたすべての日本語のノーマルツイートの長さの分布に近いことがわかる。また、図 7 から、末尾に不読符号列がある日本語のリプライの長さの分布も、Twitter に投稿されたすべての日本語のリプライの長さの分布に近いことがわかる。したがって、日本語のツイートの場合、ノーマルツイートでもリプライでも、末尾の不読符号列の有無がツイートの長さを与える影響は小さいと考えられる。同様に、図 8 と図 9 に示すように、英語のノーマルツイートとリプライの末尾の不読符号列の長さの分布は類似している。さらに図 8 から、末尾に不読符号列がある英語のノーマルツイートの長さの分布は、Twitter に投稿されたすべての英語のノーマルツイートの長さの分布に近いことがわかる。また、図 9 から、末尾に不読符号列がある英語のリプライの長さの分布も、Twitter に投稿されたすべての英語のリプライの長さの分布に近いことがわかる。したがって、英語のツイートの場合も、ノーマルツイートでもリプライでも、末尾の不読符号列の有無がツイートの長さを与える影響は小さいと考えられる。以上のことから、日本語のツイートでも英語のツイートでも、末尾の不読符号列の有無がツイートの長さを与える影響は小さいと考えられる。

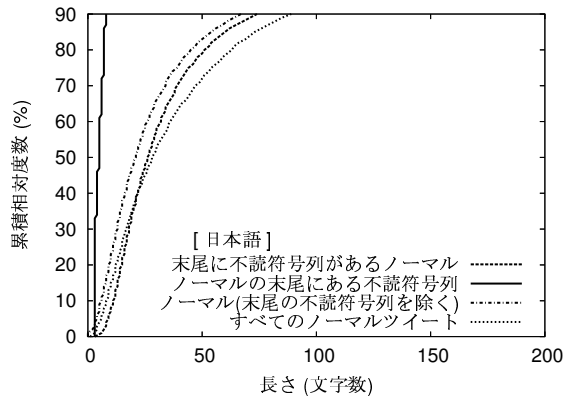


図 6 Twitter に投稿されたすべての日本語のノーマルツイート、不読符号列が末尾にある日本語のノーマルツイート、そして日本語のノーマルツイートの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

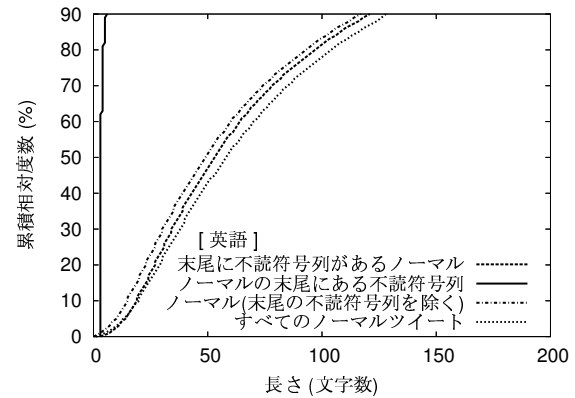


図 8 Twitter に投稿されたすべての英語のノーマルツイート、不読符号列が末尾にある英語のノーマルツイート、そして英語のノーマルツイートの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

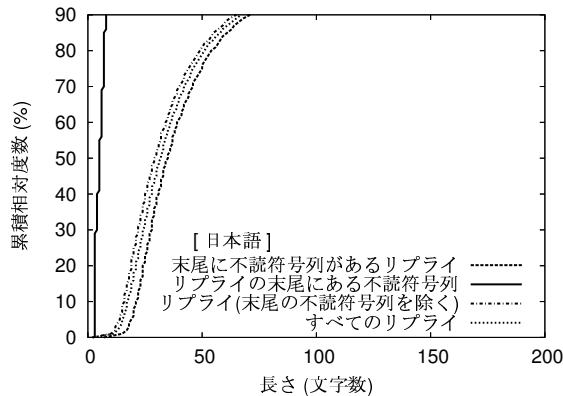


図 7 Twitter に投稿されたすべての日本語のリプライ、不読符号列が末尾にある日本語のリプライ、そして日本語のリプライの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

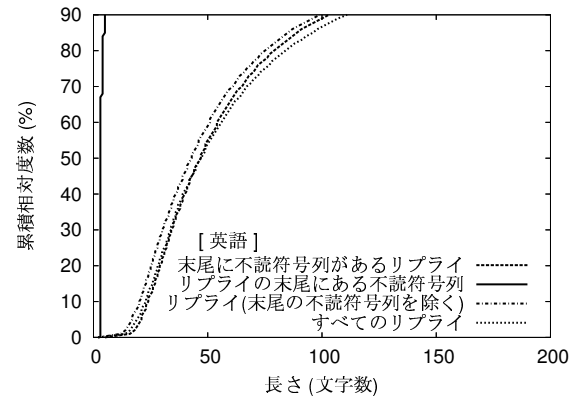


図 9 Twitter に投稿されたすべての英語のリプライ、不読符号列が末尾にある英語のリプライ、そして英語のリプライの末尾にある不読符号列とそれ以外の長さの累積相対度数分布

4. おわりに

本研究では、異なる母語話者間でも直接比較できる不読符号列を日本語と英語のツイートで比較し、末尾でそれを用いるかどうかに影響を与える要因が日本語と英語では異なることを示した。調査の対象には、2012年11月から12月までの3週間にTwitterに投稿された日本語と英語のツイートをを用いた。この日本語のツイートの13.64%、英語のツイートの10.47%の末尾で不読符号列は用いられていた。日本語のツイートの場合、末尾に不読符号列がある割合は不特定多数のユーザに宛てて投稿されたノーマルツイートでは9.4%、特定のユーザに宛てて投稿されたリプライでは17.0%と大きく異なった。一方、英語のツイートの場合、末尾に不読符号列がある割合はノーマルツイートでは7.0%、リプライでは7.6%とほとんど同じであった。したがって、言語設定で日本語を選択したユーザの場合、ツイートが特定のユーザに宛てて投稿するものであるかどうかは、ツイートの末尾で不読符号列を用いるかどうかに影響があると考えられる。一方、言語設定で英語を選択したユーザの場合、ツイートが特定のユーザに宛てて投稿する

ものであるかどうかは、ツイートの末尾で不読符号列を用いるかどうかにあまり影響がないと考えられる。これらのことから、日本語を選択したユーザにとって、特定のユーザ宛ての投稿であるかどうかは、その末尾で不読符号列を用いるかどうかに影響を与える要因の1つであると考えられる。一方、英語を選択したユーザにとってはそうではないと考えられる。

Twitterにはさまざまな言語のツイートが投稿されている。そこで日本語以外にも、ツイートが特定のユーザに宛てて投稿されたものであるかどうか、ツイートの末尾で不読符号列を用いるかどうかに影響がある言語がないか調査したい。また、日本語のツイートを対象に、特定のユーザに宛てて投稿されたツイートと不特定多数のユーザに宛てて投稿されたツイートでは、末尾で多用される不読符号列に違いがないか調査したい。

参考文献

[1] 中嶋邦裕, 梅本顕嗣, 西村涼, 渡辺靖彦, 岡田至弘, 久保圭: Yahoo!知恵袋の回答の末尾にある不読符号列の分析, 第5回知識共有コミュニティワークショップ, pp. 21-31 (2012).

- [2] 中嶋邦裕, 中山昂, 渡辺靖彦, 岡田至弘: Twitter のツイートの末尾にある不読符号列の分析, 電子情報通信学会技術研究報告, 言語理解とコミュニケーション (NLC), Vol. 112, No. 110, pp. 45–50 (2013).
- [3] Wolfson, N.: The Social Dynamics of Native and Non-native Variation in Complimenting Behavior, *The Dynamic Interlanguage: Empirical Studies in Second Language Variation*, Springer US, pp. 219–236 (1989).
- [4] Bergman, M. L. and Kasper, G.: Perception and performance in native and nonnative apology, *Interlanguage pragmatics*, Oxford University Press, pp. 82–107 (1993).
- [5] Eisenstein, M. and Bodman, J.: Expressing gratitude in American English, *Interlanguage pragmatics*, Oxford University Press, pp. 64–81 (1993).
- [6] Tanaka, S. and Kawade, S.: Politeness Strategies and Second Language Acquisition, *Studies in Second Language Acquisition*, Vol. 5, No. 1, pp. 18–33 (1982).
- [7] Beebe, L. M., Takahashi, T. and Uliss-Weltz, R.: Pragmatic transfer in ESL refusals, *Developing communicative competence in a second language*, Newbury House Publishers, pp. 55–73 (1990).
- [8] Deschamps, A.: The syntactical distribution of pauses in English spoken as a second language by French students, *Temporal variables in speech*, De Gruyter Mouton Publishers, pp. 255–262 (1980).
- [9] Ishizaki, A.: How Does a Learner Leave a Pause When Reading Japanese Aloud?: A Comparison of English, French, Chinese and Korean Learners of Japanese and Native Japanese Speakers (in Japanese), *Japanese-Language Education around the Globe*, Vol. 15, pp. 75–89 (2005).
- [10] 堀口純子: 上級日本語学習者の対話における聞き手としての言語行動, 日本語教育, Vol. 71, pp. 16–32 (1990).
- [11] LoCastro, V.: Aizuchi: A Japanese conversational routine, *Discourse Across Cultures: Strategies in World Englishes*, Prentice Hall, pp. 101–113 (1987).
- [12] Tera, A., Shirai, K., Yuizono, T. and Sugiyama, K.: Analysis of Eye Movements and Linguistic Boundaries in a Text for the Investigation of Japanese Reading Processes, *IEICE Transactions on Information and Systems*, Vol. E91.D, No. 11, pp. 2560–2567 (2008).
- [13] Minematsu, N.: Pronunciation assessment based upon the phonological distortions observed in language learners’ utterances, *Proc. of 8th International Conference on Spoken Language Processing (ICSLP 2004)*, pp. 1669–1672 (2004).
- [14] Tsubota, Y., Kawahara, T. and Dantsuji, M.: Practical Use of English Pronunciation System for Japanese Students in the CALL Classroom, *Proc. of INTERSPEECH 2004*, pp. 1689–1692 (2004).
- [15] Nakagawa, S., Nakamura, N. and Mori, K.: A Statistical Method of Evaluating Pronunciation Proficiency for English Words Spoken by Japanese (Speech and Hearing), *IEICE Transactions on Information and Systems*, Vol. E87.D, No. 7, pp. 1917–1922 (2004).
- [16] Raux, A. and Kawahara, T.: Automatic Intelligibility Assessment and Diagnosis of Critical Pronunciation Errors for Computer-Assisted Pronunciation Learning, *Proc. of 7th International Conference on Spoken Language Processing (ICSLP 2002)*, pp. 737–740 (2002).
- [17] Muto, M., Sagisaka, Y., Naito, T., Maeki, D., Kondo, A. and Shirai, K.: Corpus-based modeling of naturalness estimation in timing control for non-native speech, *Proc. of EUROSPEECH 2003*, pp. 401–404 (2003).
- [18] Nakamura, S., Tsubaki, H., Kondo, Y., Nakano, M. and Sagisaka, Y.: Tempo-normalized measurement and test set dependency in objective evaluation of english learners’ timing characteristics, *Proc. International Congress of Phonetic Sciences (ICPhS)*, pp. 1733–1736 (2007).
- [19] Ito, A., Nagasawa, T., Ogasawara, H., Suzuki, M. and Makino, S.: Automatic Detection of English Mispronunciation Using Speaker Adaptation and Automatic Assessment of English Intonation and Rhythm, *Educational technology research*, Vol. 29, pp. 13–23 (2006).
- [20] Maynard, S. K.: On back-channel behavior in Japanese and English casual conversation, *Linguistics*, Vol. 24, No. 6, pp. 1079–1108 (1986).
- [21] Miller, L.: Verbal listening behavior in conversations between Japanese and Americans, *The Pragmatics of Intercultural and International Communication*, John Benjamins Publishing, pp. 110–130 (1991).
- [22] White, S.: Backchannels across cultures: A study of Americans and Japanese, *Language in Society*, Vol. 18, No. 1, pp. 59–76 (1989).
- [23] 陳姿菁: 日本語におけるあいづち研究の概観及びその展望, 言語文化と日本語教育, 増刊特集号 第二言語習得・教育の研究最前線, Vol. 2002, pp. 222–235 (2002).
- [24] Kato, S., Kato, Y., Shimamine, Y. and Yanagisawa, M.: Analysis of functions of emoticons in e-mail communication by mobile phone: Investigation of effects of degrees of intimacy with partners, *the journal of Japan Society of Educational Information*, Vol. 24, No. 2, pp. 47–55 (2008).