

人工無脳は無能でない

—人工無脳に基づいた英語ライティング学習支援の効果測定—

橋口 友哉[†] 岸本 優也[†] 河原源太郎[†] 松本 佳樹[†] Driss Sadoun^{††}

永田 亮[†]

[†] 甲南大学 ^{††} PostLab

E-mail: †nagata-nlp2019@hyogo-u.ac.jp.

1. はじめに

英文ライティングの学習において、たくさん書くことは重要な学習活動であるが、学習者にとって困難が伴うことが少なくない。特に、初級から中級の学習者にとっては、たくさん書くことは敷居が高い。そもそも何を書くべきか思いつかないという問題がある。仮に、思いついたとしても、それをどのように英語で表現したらよいか分からないという問題もある。学習者が知っている単語でも、文章中で適切に使うことは難しいことが少なくない。このような状況で、モチベーションを保ちつつ、ライティング学習を続けるというのは多大な労力を要する。教師などからのフィードバックがなくとも、書き続けることでライティングの正確さが向上するという報告 [1] もあり、学習者にいかにたくさん書かせるかということは、ライティング学習が成功するための一つの大きな鍵といえる。

以上のような背景を踏まえ、本稿では、次のような仮説を立て、新しい英文ライティング学習支援方法を提案する：

仮説： 単純なチャットボット（人工無脳）でも、たくさん書くという学習活動を支援することには効果がある。

この仮説の根拠は次のようなものである。チャットボットが学習者に対して行う対話や質問により、学習者は書く内容を思い浮かべやすくなることが予想される。すなわち、チャットボットとのやり取りを通じて、より幅広い内容についてライティングを行うと期待できる。更に、チャットボットの発話中の単語や表現を、学習者の英文に取り入れるという効果もみられる可能性がある。英語母語話者との会話で、新しい単語や表現を学び、更にその会話中で使用するということは実際によく行われることである。また、学習者が一人で淡々と書くよりも、相槌や質問といったチャットボットからのリアクションがあるほうが書くためのモチベーションが保ちやすいということも予想できる。

本研究では、この仮説の検証を行った。具体的には、チャットボットを用いてライティングを行った場合とそうでない場合で、書く量にどの程度差異が生まれるかを測定した。本研究で用いたチャットボットは ELIZA [2] である。ELIZA のス

クリプトを、ライティングのトピックに沿った発話や質問となるように設計した。学習者は ELIZA と対話を行い、その結果をライティング結果とした。一方で、チャットボットを用いない場合は、Web ブラウザ上で動作するエディタで学習者はライティングを行った。これらのライティングの結果から、両者の書く量にどの程度差異が生まれるかを測定した。更に、得られたデータを定量的および定性的に分析し、両者の違いを考察した。仮に、ELIZA のような単純なアルゴリズムに基づいたチャットボットでも、たくさん書かせることに効果があるとすれば、ライティング学習支援の新たな可能性を示唆することになる。そのような結果は、より高度な対話システムを用いた学習支援システムの開発に繋がる。

2. 仮説の検証方法

1. の仮説を検証するため、学習者に実際にライティングを行ってもらった実験を行った。具体的には、チャットボットを利用したライティング（以降、チャットボットありと表記する）と用いないライティング（以降、チャットボットなしと表記する）で、書く量にどの程度差異が生まれるかを測定した。

既に述べたように、用いたチャットボットは ELIZA に基づく。実装したシステムのサンプル画面を図 1 に示す。図 1 の左側のアイコンから出ている吹き出しのテキストがチャットボットの発話である。この図のように、まずはチャットボットが学習者に話しかけることでライティングが始まる。学習者は、チャットボットの発話に対して返答するという形でライティングは進んでいく。その結果は、ログファイルとして保存される。

本検証実験で用いた ELIZA のスクリプトは、英語教員と言語処理技術者が協力して作成した。ライティングのトピックは、「自己紹介」と「趣味」に関するもの二種類とした。それぞれのトピックに対して、ライティングが続くようという観点からスクリプトを作成した。スクリプト中のキーワードについては「自己紹介」が 30 個、「趣味」が 28 個のスクリプトとなった。また、キーワードに対する返答については「自己紹介」が 83 種類、「趣味」が 79 種類となった。

一方、チャットボットなしのライティングでは、Web ブラウザ上で操作を行うエディタとした（図 2）。このエディタ

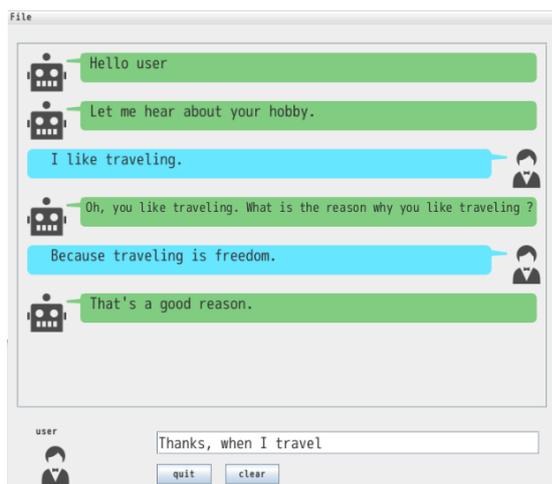


図 1: チャットボットを利用した場合のライティング画面.

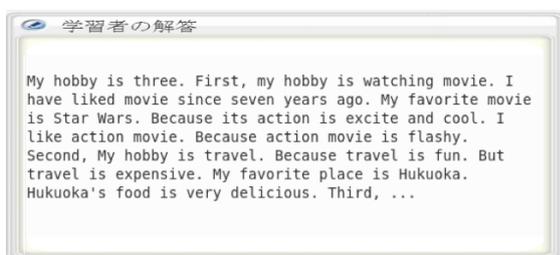


図 2: チャットボットなしの場合 (通常のエディタ) のライティング画面.

は、文章の入力や編集という基本的な操作が可能である。したがって、通常のライティングのようにエディタ上にテキストを入力することで学習者はライティングを行う。

以上の二種類のシステムを用いて、学習者に英文を書いてもらった。学習者は大学生 18 名である。各学習者は、チャットボットあり/なしの各一回、計二回のライティングを行った。各回 50 分のライティング時間とした。ライティングの際には、辞書などは用いず、学習者は自力でライティングを行った^(注 1)。

本検証実験には、チャットボットあり/なしと二種類のライティングトピックがあるため、計四種類の条件があることになる。チャットボットの使用順とトピックが書く量に与える影響を極力減らすため、18 名の学習者をランダムかつ均等に四種類の条件に振り分けた。ただし、18 名で完全に均等にはできないため、チャットボットを先に使用して「自己紹介」についてライティングという条件を 5 名、通常のエディタを先に使用して「自己紹介」についてライティングという条件を 5 名、残りの条件を各 4 名と振り分けた。

(注 1) : ただし、チャットボットの発話を参考にすることは除く。

検証実験は二回にわけて行った。一回目は 2016 年に学習者 10 名で行った。二回目は 2017 年に別の 9 名で行った。ただし、内 1 名については正しくデータを取得できなかったため、最終的に得られたデータは 8 名分である。したがって、計 18 名の延べ 36 回分のライティングデータを得た。なお、ライティングは一日に一回とした (すなわち、各学習者は二つのトピックについて別の日にライティングを行った)。

2017 年に実施した検証実験では、ライティング終了後、学習者に自由記述式で各システムについての感想を書いてもらった。チャットボットありとなし、それぞれ 8 人分の計 16 のアンケート結果を得た。

検証実験の結果得られたライティングデータの量を計数するために、次の前処理を施した。チャットボットありの条件では、チャットボットが同じ質問を繰り返し、それに対して学習者も同じ回答を繰り返すことがある。このような同一内容をそのまま計数することは、どの程度書く量に差異があるかを検証する上で好ましくない。そこで、同じ質問に対して同じ回答をしている部分は削除した。また、ほぼ同じ表層で、同一内容を表している場合も削除した (この場合、より長い方を残した)。この作業は、目視で確認し人手で行った。また、実質的なライティング量を計数するため、次の 14 種類の記号が単独で出現した場合は削除した: 「」, 「,」, 「!」, 「?」, 「“」, 「”」, 「'」, 「'」, 「/」, 「(」, 「)」, 「...」, 「-」, 「:」。なお、単語の同定には、Stanford-corenlp^(注 2)を用いた。

3. 検証結果

測定結果を表 1 に示す。「単語トークン数平均値」とは、学習者が書いた単語の総数をチャットボットあり/なしごとに平均した値である。同様に、「単語タイプ数平均値」とは、学習者が書いた単語の種類数をチャットボットあり/なしごとに平均した値である。

表 1 から、チャットボットありの条件では、チャットボットが同じ質問を繰り返して行うことがあるので、単純に単語トークン数を数えるとライティング量を大きく見積もってしまうことがわかる。今回の検証手順のように、同一質問に対する同一内容の返答かを目視によりチェックして、実質のライティング量を測定する必要があるといえる。

同一内容を削除した場合、チャットボットありでは、チャットボットなしと比べて、単語トークン数の平均値は約 1.2 倍となった。より詳細に両者の違いを分析するために、チャットボットなし/ありを x/y 軸にとった単語トークン数の散布図を図 3 に示す。各プロットは学習者一人ひとりに対応する。同図からわかるように、大部分の学習者において、チャットボットを利用した場合、通常のエディタを用いたときの同等以上の量を書いたことがわかる。ただし、両条件の差につ

(注 2) : <https://stanfordnlp.github.io/CoreNLP/>

表 1: チャットボットあり／なしのライティング量の平均値.

	単語トークン数平均値	単語タイプ数平均値
チャットボットあり	295	117
チャットボットあり (同一内容を削除)	247	115
チャットボットなし	204	105

いて、 t 検定を行ったところ、 p 値は 0.064 となり、有意水準 5% では有意とはいえないという結果となった。しかしながら、 p 値は比較的小さな値を示しておりチャットボットを用いることで学習者はたくさん書ける傾向にあることを示唆する。

一方、単語タイプ数の差 (注 3) についても t 検定を行ったところ p 値は 0.270 とより大きな値で有意とはいえないという結果となった。今回用いたチャットボットは、より多くの種類の単語を学習者に使わせるように設計されたわけではないので、この結果はある意味妥当であるといえる。ただし、4. で説明するように、一部の学習者については、チャットボットの発話に出現した単語をライティングに利用していると思われる事例が確認された。

また、アンケートの結果は次のように要約される。チャットボットなしの条件では、「単語を思い出せなくて書くのに苦労した。」(3 名)、「辞書を使って書きたかった。」(1 名) という、自力で書くことの難しさに関するものが多数を占めた (なお、アンケートの全回答数はチャットボットあり／なし、各 8 名分である)。

一方、チャットボットありの条件では、「対話形式の方が楽しくて時間が早く感じた。」(1 名)、「対話形式の方が話題にあわせて言葉を引き出してくれようとするので楽だった。」(1 名)、「英語を書く話題に悩まなかった。」(1 名) のように、対話形式でのライティングに対してポジティブな意見がみられた。しかしながら、同時に、「同じ質問が繰り返されて返答に困った。」(4 名)、「システムの質問に具体性がなく返答しにくい。」(3 名) というネガティブな意見も多く見られた。

4. 考 察

3. の結果から、学習者はチャットボットを利用することでたくさん書ける傾向にあることがわかった。本節では、実験結果を更に詳細に分析することで、チャットボットをライティングに用いることのメリットとデメリットを考察する。

アンケートの結果から、チャットボットが繰り返し同じ質問をすることを敬遠する学習者がいることが明らかになった。現行のチャットボットでは、アルゴリズムの性質上、長時間使用すると同じ質問を繰り返すことは避けられない。この点

(注 3) : こちらも同一内容を削除した場合を対象にしてである。チャットボットあり／なしの平均値の比は約 1.1 である。

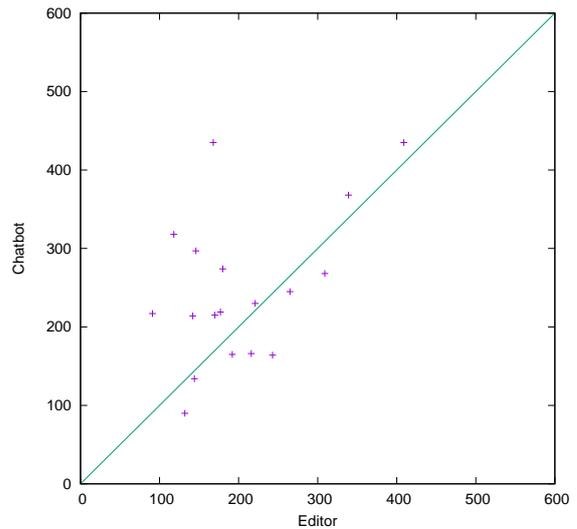


図 3: チャットボットあり／なしにおける単語トークン数の散布図

については、一度した質問はしないなどの制御が必要となる。

ただし、ライティング結果を実際に分析したところ、同じ質問を二度することに良い面もあることが確認できた。同じ質問が繰り返された際に、以前の内容を振り返り、二度目のライティングでは、自発的に間違いを修正したり、表現を工夫したりしている例が観測された。全てのデータを分析したところ、同じ質問により文の質が向上した学習者は 18 名中 10 名であった。具体的には、綴り誤りの訂正、主語動詞の一致誤りの訂正、冠詞の抜けの訂正、情報の追加などが確認できた。以上のことは、同じような質問を適度に繰り返すことにはメリットもあることを示唆する。

更に、チャットボットありの条件のみで使用された単語について調査を行った。表 2 に、チャットボットありの条件のみで使用された単語とその頻度を示す。表 2 より、「lyrics」、「elder」、「active」など比較的難易度が高い単語が使用されていることがわかる。

実際に対話の流れを分析してみると、チャットボットの発話にこれらの単語が出現し、その単語を学習者が自身の次の返答で使用している例が多数確認できた。表 3 は、表 2 の単語のうち、チャットボットの発話に出現したもの、かつ、その発話の直後に、学習者が自身の発話でその単語を使用した割合を示す。表 3 から、チャットボットありの条件のみで使用された単語のうち約半分は、チャットボットの発話から流用したことがわかる。このことから、チャットボットの発話に (すなわち、スクリプトに)、学習させたい単語を入れておくことで、効率的に単語の学習が行われる可能性があることを示唆する。

一方、チャットボットをライティング学習に用いることのデメリットの一つに文のつながりや文章の流れといった文章

表 2: チャットボットありの条件のみで使用された単語.

単語	頻度
yes	29
lyrics	18
likes	15
born	13
nickname	12
yesterday	12
an	8
n't	7
talking	7
remember	6
active	5
elder	4
sports	4

表 3: チャットボットの発話に出現した単語を使用した学習者の割合 (%).

単語	単語を使用した学習者の割合
born	100
nickname	89
active	56
an	56
lyrics	56
elder	44
sports	44

構成に関する検討が行われにくいということがある。なぜなら、ELIZA に基づいたチャットボットは、発話の選択にランダム性を伴うためである。そのような発話に対して学習者は自身の返答を書くため、文章構成について検討が行われにくいのは、ある意味当然である。一つの改善方法としては、チャットボットに文章の流れを考慮するような制御機構を導入することが考えられる。しかしながら、そのような制御機構を導入することは容易ではない。また、仮に導入できたとしても、学習者が明示的に文章構成を考えてライティングを行うとは限らない（文章の流れは考えず、学習者はチャットボットの発話に単に返答するということが予想される）。

以上を踏まえ、本稿では、チャットボットを用いた、次のようなライティング学習活動を提案する。まず、今回の検証実験のように、学習者はチャットボットを利用してライティングを行う。チャットボットの利用により、学習者は、より幅広い内容について、たくさんの量を書くことが期待できる。一方、上で指摘したように、ここで得られたライティング結果は、文章構成についてはほとんど検討されていないものである可能性が高い。そこで、学習者は、チャットボットの発話と自分の発話（ライティング）を振り返り、内容の取捨選択、グルーピング、書く順番などを検討する。すなわち、こ

の時点で文章構成を検討する。加えて、誤りの修正、既書いた内容の掘り下げ、新たな内容の追加などの検討をしてもよい。最後に、この結果に基づいて、通常のライティングを行うという方法である。以上の過程により、文章構成も考慮しつつ、より幅広い内容についてたくさん書くという学習活動を行うことができる。

5. おわりに

本稿では、「単純なチャットボット（人工無脳）でも、たくさん書くという学習活動を支援することには効果がある。」という仮説を検証した。具体的には、チャットボットを用いたライティングと通常のライティングとでどの程度書く量に差が出るかを測定した。その結果、チャットボットを用いるとより多く書ける傾向にあることが確かめられた（チャットボットありの場合は 1.2 倍多く書ける結果となった。ただし、 p 値は 0.064 である）。更に、得られたデータを定量的および定性的に分析した。その結果、チャットボットを利用することで学習者は、(1) 自発的に、間違いの修正や表現の改善を行うことがある、(2) チャットボットの発話に出現した単語を自身のライティングに使用する、ということも確認できた。以上の結果と分析に基づき、チャットボットを用いたライティングと従来のライティングとを組み合わせた新しいライティング学習支援方法を提案した。

参考文献

- [1] R. Nagata and K. Nakatani, "Evaluating performance of grammatical error detection to maximize learning effect," Proc. of 23rd International Conference on Computational Linguistics, poster volume, pp.894-900, 2010.
- [2] J. Weizenbaum, "ELIZA — a computer program for the study of natural language communication between man and machine," Communications of the ACM, vol.9, no.1, pp.36-45, 2000.