

例文引用をベースとした英文作成ツール

嶋屋 友佳¹ 松野 祐典¹ 宮下 芳明^{1,2}

概要: 本稿では、ユーザが手軽にデータベースの例文を引用し、文章として構築するための英文作成ツールを提案する。提案ツールではまず、意図に近い例文や表現を、キーボードによる日本語入力で検索する。そして、例文の単語や熟語などをマウス操作で移動・編集し、レイアウトしながら英文を作成していく。これまでの英文作成は、辞書サイト等で単語や例文の検索を行い、日本語入力の ON/OFF を切り替えながら複数のアプリケーションやウィンドウをまたがってコピー&ペーストする作業になるか、あるいは、完全な機械翻訳で済ませ、翻訳された英文にはほとんど手を加えない（あるいは見ない）、という両極端なものになりがちであった。提案ツールを用いることで、文法的には不適格かもしれないが「伝わる英文」を作成でき、会話調で行う短文コミュニケーションサービスで有効であると考えられる。

English Writing Tool Based on Sample Sentences

YUKA SHIMAYA¹ YUSUKE MATSUNO¹ HOMEI MIYASHITA^{1,2}

Abstract: In this paper, we propose an English writing tool to easily take sample sentences from a database and construct words into phrases. The user first types Japanese words on a keyboard and looks up English sample sentences or expressions that are close in meanings of what is input. Then the user moves and edits the words or phrases with a mouse, lay them out and construct an English text. Until now, writing English text/phrases process had two variations. One uses a dictionary database to search for words or sample sentences among multiple applications and windows, and then copy and paste words while turning on and off the Japanese input function. The other uses machine translation and the user never see the text result. Our proposed tool may not necessarily result in grammatically well-formed sentence, however the user can write understandable English text. We considered that it is effective as a microblogging site in conversational-style communication.

1. はじめに

1.1 背景

Twitter[1] のような同期型の短文コミュニケーションサービスが世界的に広がり、国や地域を問わない交流の機会が増えた。しかし、このような短い文章でのグローバルなやりとりの障害となるのが言語である。例えば Twitter のようなタイムライン表示形式を持つサービスにおいて、英語圏フォロワーの一つのツイートに対しリプライをしよ

うとした際、英語知識の少ない書き手が用いる手段は主に二つある。一つは、機械翻訳に頼る方法である。しかし、機械翻訳によって得た英文の内容が相手に伝わるかどうかは、利用した翻訳サービスの精度に依存し、その精度もあまり高くないため、内容が通じなかったり、誤解を招いてしまったりすることも多い。即時的なコミュニケーションの場合は特に、正しいニュアンスで出力された英文なのかをあまり確認せずに返信してしまうこともある。もう一つは、オンライン辞書サイトを利用し、キーボードによる日本語入力の検索とマウス操作によるコピー&ペースト、キーボードによる直接入力での文法訂正を繰り返して作成する方法である。このような作業はキーボード入力やマウス操作、アプリケーションやウィンドウの切り替えなどの動作を伴うため効率がよいとはいえない。いざ完成した英文を

¹ 明治大学大学院 理工学研究科 新領域創造専攻 デジタルコンテンツ系

Program in Digital Contents Studies, Programs in Frontier Science and Innovation, Graduate School of Science and Technology, Meiji University

² 科学技術振興機構 CREST
JST CREST

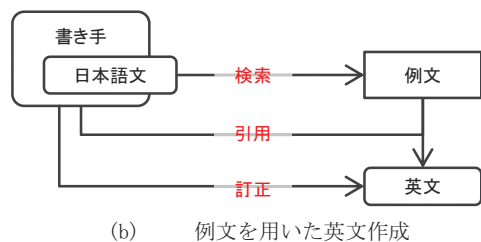
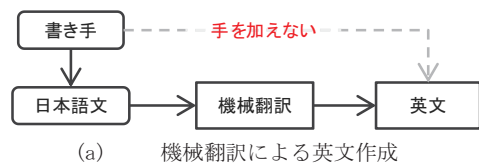
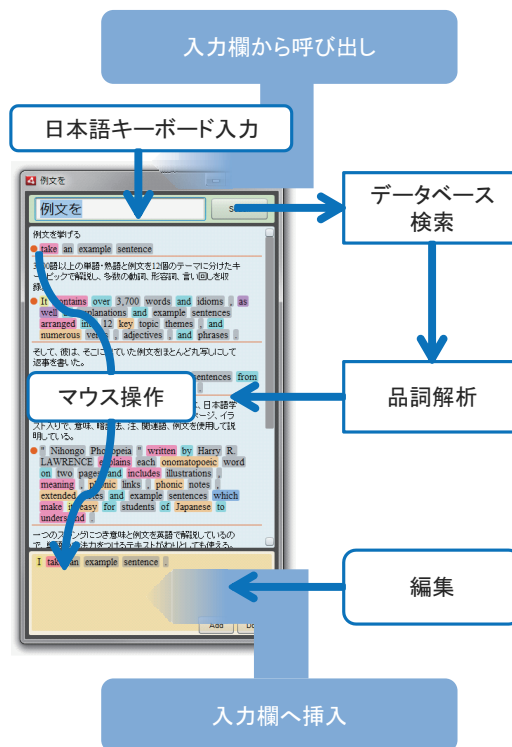


図 1 英文作成
Fig. 1 English writing



リプライしようとしても、既にリプライ先のツイートが流れてしまうということもある。また、リアルタイムに近いコミュニケーションでは、文章作成に時間がかかると返信するタイミングを失うことがある。このように、機械翻訳に頼れば短時間で作成できるが内容が不正確になりがちとなり、辞書サイト等を駆使した作文では内容が正確になるが今度は時間がかかってしまう。短文コミュニケーションにおいて重要なことは、ある程度の正確さをもった文章を、ある程度の速度で作成できることではあるはずだが、どちらの方法もそれに適していないということになる。

そこで本稿では、ユーザが手軽に辞書データベースの例文を引用し、目的とする文章を構築するための他言語文章作成ツールを提案する。今回はインターネット上で最も多く利用されている言語である英語 [2] を対象としたツールとして実装した。提案する英文作成ツールではまず、ユーザが作成したい英文の意図に近い例文や表現を、キーボードによる日本語入力で検索する。そして、例文の単語や熟語などをマウス操作で移動・編集し、レイアウトしながら英文を作成していくものとした。

1.2 例文を用いた英文作成

例文を用いた英文作成は、機械翻訳と直接入力による英文作成の中間に位置する手法であるといえる (図 1)。機械翻訳による英文作成では書き手の英語文法・表現ともに知識が乏しい場合、入力した日本語文に対して自動的に出力された英文に手を加えないことがある。そのため、書き手と直接的な繋がりがなく、意図しない英文となる場合がある。一方、直接入力による英文作成は書き手の持つ英語

知識を用いて自ら文章として構築するため、書き手の意思が直接的に反映された英文となる。例文を用いた英文作成は、英語知識を辞書データベースからの自動的な検索結果に頼りつつ、引用と訂正を書き手自らが行う。それにより、直接入力と同様に書き手の意思を含んだ英文を作成することができると思われる。

1.3 短文コミュニケーション

インターネット上における短い文章を用いたコミュニケーションの特徴は、情報量が少なく文法構造などをあまり考慮しない会話調の文章、すなわち文法的不適格文であることが多い。このようなコミュニケーションにおいては、言語が異なるとしても、相手に意図の伝わる文章であればよいのではないかと筆者らは考える。たとえ、書き手の文章に多少の文法の間違ひがあっても、読み手は好意的に解釈するからである。提案ツールによって作成される英文は、文法的には不適格かもしれないが「伝わる英文」であり、会話調で行う短文コミュニケーションサービスで有効であると考えられる。

2. 提案ツール

2.1 概要

提案する英文作成ツールでは、図 2 のように文章作成時にショートカット機能を用いることで別ウィンドウが開き、日本語の語句を入力とした例文・表現の検索と例文を用いた編集作業を同ウィンドウ上で行えるものとする。そ

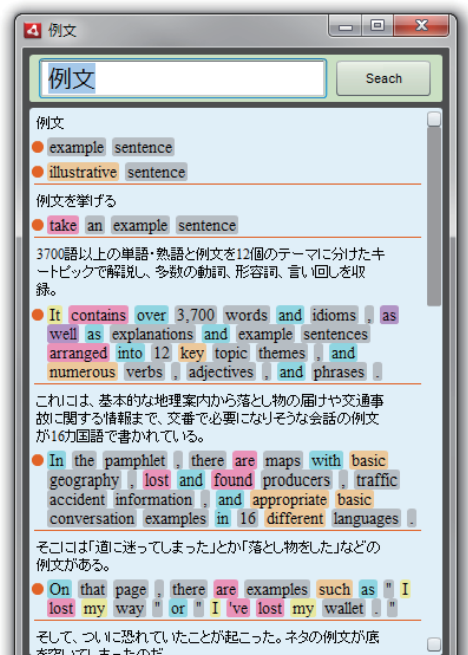
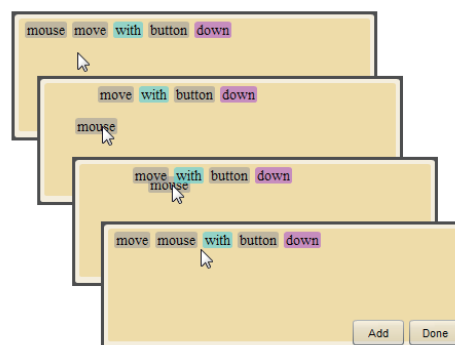
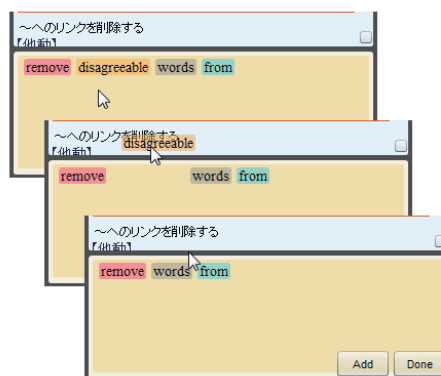


図 3 検索結果の一覧

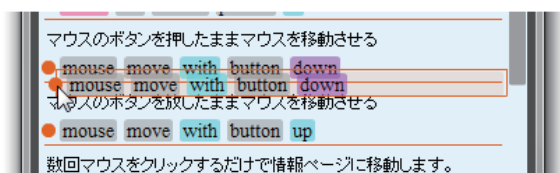
Fig. 3 List of search result



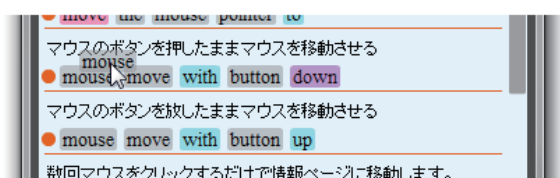
(a) 単語オブジェクトの移動



(b) 単語オブジェクトの削除



(a) 英文の全単語オブジェクト移動



(b) 英文内の一単語オブジェクト移動

図 4 単語オブジェクトの移動

Fig. 4 Move of word object

れにより、検索と編集をアプリケーションやウィンドウの切り替えなどの動作を行う必要がなくなり手間を省くことができる。また、キーボード入力により例文を検索し、単語の移動・編集、レイアウトをマウスのドラッグ・クリックなどの操作によって行う。単語をオブジェクトとして扱うことで文章の構築を GUI とし、よりマウスの操作性を考慮したものとした。最終的な出力時に、文章入力を行っていた位置へ作成された英文を挿入するという手法をとることにより、英文作成のみを独立して行うことができる。

図 5 編集領域内での移動・削除

Fig. 5 Move and delete in edit area

2.2 操作

2.2.1 例文の検索

検索窓に作成したいと考える英文に関する日本語の語句をキーボードにより入力する。複数語句による AND 検索を行うこともできる。検索結果は図 3 のように、その語句を含む日本語文と対応する英文を例文として辞書のデータベースから取得し、再構築し一覧表示される。日本語文を読み、検索結果一覧から相応しいと思われる例文を探し出す。品詞ごとに色分けされた単語オブジェクトを見て、文法の確認をする。例文の検索は編集操作中、随時行うことができる。

2.2.2 単語オブジェクトの選択

検索結果一覧から編集領域への単語オブジェクト移動には図 4 のように、

(a) 英文の全単語オブジェクト移動

(b) 英文内の一単語オブジェクト移動

の二つの方法を用いることとする。(a) は英文の文頭に付加されている丸いアイコンを、(b) は個々の単語オブジェクトをマウスでドラッグし、編集領域にドロップすることで移動する。主にはじめのベースとなる例文を (a)、組みかけ文内に単語オブジェクトを挿入する場合に (b) の手法を用いて移動する。

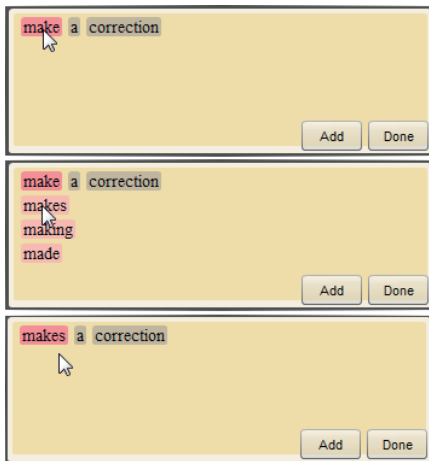


図 6 単語の訂正

Fig. 6 Correcting words

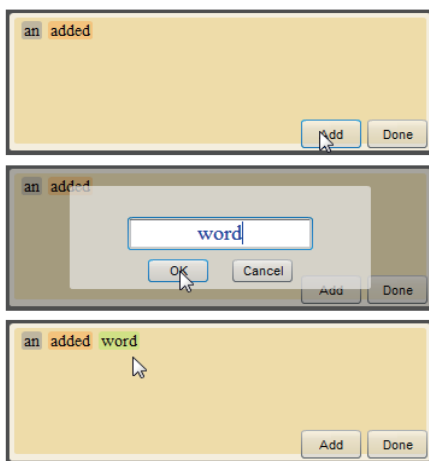


図 7 単語オブジェクトの追加

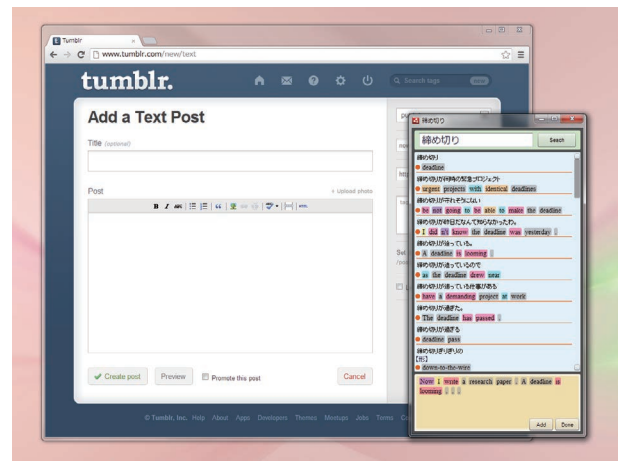
Fig. 7 Add word object

2.2.3 編集領域

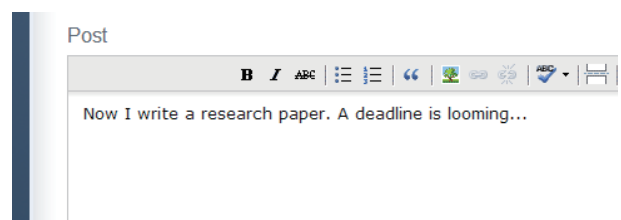
編集領域では、マウス操作による単語オブジェクトの移動・削除、人称・時制の修正などを行う。検索結果の一覧からの編集領域への移動操作と同様に、単語オブジェクトをドラッグ、ドロップすることで移動、任意の位置へ挿入できる (図 5(a))。その際、編集領域以外の場所でドロップされた場合、その単語オブジェクトは組み合わせ文から削除される (図 5(b))。組み合わせ文の単語オブジェクトは、各操作の際には自動整列され等間隔に並べられる。

図 6 のように人称や動詞、形容詞・副詞などの単語オブジェクトについては、マウスの右クリックによりプルダウンメニューが開く。プルダウンメニューには、該当する単語オブジェクトに対してデータベースから取得した活用・変化形が表示される。マウスで適切なものを選択することで、キーボード入力なしに目的の形に訂正できる。

固有名詞等にはキーボード入力を行う必要がある。そのために、単語オブジェクト生成機能を用意している。図 7 のように編集領域に用意された追加ボタンをクリックす



(a) テキスト入力欄から呼び出して使用



(b) 挿入した英文

図 8 英文作成例

Fig. 8 Example of English writing

ることでキーボード入力が可能となる。単語を入力し確定後、新たに入力された単語オブジェクトが組み合わせ文の末尾に挿入される。

3. 使用例

図 8 は実装した英文作成ツールを実際に使用し、英文を組み立て、挿入した例である。ユーザは、「論文執筆中で締め切りが近い」という意図の内容をインターネット上で交流のある人々に英語で伝えるためにツールを使用した。テキスト入力欄からショートカットキーにより呼び出したツールで、はじめに「論文 書」で検索した。例文の中から「研究論文を書く」という日本語文の英訳である“write a research paper”を編集領域へマウス操作で英文の全単語オブジェクト移動した。次に「現在」という語句で検索した。「現在」を意味する英単語は複数得られる。そこで、例文の中で作成したい文章と同様の意味合いで用いられているものを各例文をそれぞれ比較して選ぶ。その結果として“Now”を文頭に含んだ英文から一単語オブジェクトを選び編集領域へ移動、組み合わせ文の文頭へ挿入した。同様に検索結果一覧にある“I”を一単語オブジェクトとして移動、“Now”と“write”の間に挿入した。次に「締め切り」という単語で検索した。結果の中の「締め切りが迫っている。」という日本語文の英訳である“A deadline is looming.”を編集領域へマウス操作で英文の全単語オブジェクト移動した。エモーション系の辞書データは用意されていない

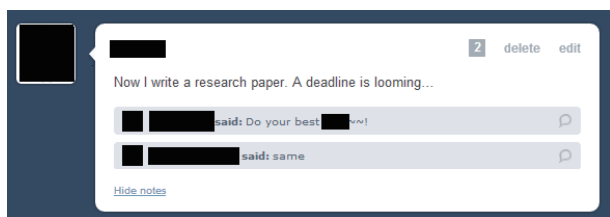


図 9 投稿に対する反応
Fig. 9 Replies to post

ため、ピリオドを2つ追加挿入することで憂鬱、というような状態を表現した。組み立てた文は、Done ボタンによりテキスト入力欄へ挿入され、通常のテキスト入力と同じ状態の文となった。完成した英文である“Now I write a research paper. A deadline is looming...”は、「今、私は論文を書いています。締め切りが近い…。」という日本語の意図を持つとユーザが考えて組み立てたものである。この例では、組みかけ文内の単語オブジェクトに対する編集操作を行うことはなく、実際の英文作成にかかった時間は数分であった。

実際に作成した英文を、ミニブログであり SNS サービスの一つである Tumblr[3] に投稿したところ、二人のフォロワーからリプライのリアクションがあった。フォロワーには事前にツールを使用することは伝えていない。また、投稿者のフォロワーはほぼ日本人以外のインターネット上で主に英語を用いてコミュニケーションをしている人々である。図 9 はフォロワーからのリプライを表示したものである。投稿した英文に対するフォロワー一人目からのリプライ (Tumblr の投稿したアイテムに対するリアクションは下から上へ時系列表記されている) は“same(同じ)”のように同様な状態であるという同情を示すものであった。二人目からのリプライは“Do your best, ■■■~! (頑張って, ■■■~~!)”のように投稿者を励ますものであった。これらの反応から、投稿者の意図した内容が正しく伝わったことが分かる。

4. 関連研究

本研究に対して、主に英文理解・作成と、入力・作文のためのツールやシステムの二方向の関連する既存研究が挙げられる。五十嵐は、不完全な機械翻訳システムを利用しつつ、より円滑な異言語間コミュニケーションを実現するためのインタラクティブなテキスト編集・表示システム [4] を提案している。書き手が文の構造を明示的に指定し、その構造をそのまま読み手に提示することで読解支援を行っている。松原らの用例文検索システム ESCORT[5] は英文作成の支援を目的として開発され、キーワード系列を入力とし、キーワード間に構造的な関係が存在する英文を提示する。武田らの英文書作成支援システム [6] では、データベースの中にある例文を使い、修正や合成、編集を繰り返

すことにより、目的の文章を作成する。意味処理を行わずに含意検索を実現し、学習機能により検索を繰り返すほど効率上がる結果を得ている。大鹿らは、検索エンジンを用いて訳語候補を自動的に提示できる英作文支援システム [7] を提案している。フレーズ検索やワイルドカードを用いた検討、多義語の検討を行うことができる。

池田らは、利用したい言語に不慣れな言語学習途上のユーザでも簡単に Web 上にある所望の言語の情報資源に検索エンジンを用いてアクセスできる、多言語に対応した入力支援システム [8] を提案している。綾塚らは、短い文の入力をターゲットとし、手数を予測変換よりも少なく抑え、文のバリエーションをなるべく多くする作文手法「完成文変更方式」を提案している。試作システム GhostTweet[9] では、ユーザが一から文を作成するのではなく、最初に自動的に生成された候補の文が提示され、選択することで Twitter にツイートできる。

キーボードとペンという2つの異なる操作手法を用いたアプリケーションとして辰川らのノートエディタ [10] が挙げられる。ノートエディタではテキスト入力をキーボード、図形の描画をペンと分けている。それにより、パソコンと手書きそれぞれ読みやすさ、編集と検索のしやすさ、図形を容易に描画できるというメリットを併せ持つのである。またビジュアルプログラミング言語においても、命令要素をマウスのドラッグ操作により、組み立てて編集するシステムが存在する。プログラミン [11] や M. Resnick らの Scratch[12] は子供が GUI ベースのプログラミング環境により、自分で描いた絵を動かすことができる。Alice[13] では、学生が 3D オブジェクトを変更してアニメーションを生成するプログラムを書くことができる。マウスのクリック操作により、命令要素を作成し、ドラッグ操作で条件分岐を追加できる。

文章の訂正機能としてマウス操作を用いるインタフェースが存在する。緒方らの音声訂正インタフェース [14] や PodCastle[15] における編集機能は、音声認識結果の単語に加え、認識処理過程において可能性の高い競合する候補をリストとして保持している。ユーザは認識誤りを発見した際、そのリストの中から正しい単語をクリックし選択することができる。正しい単語がない場合は、キーボードを用いて自由にタイプすることもできる。Google 翻訳 [16] においても、機械翻訳によって訳された文章中の単語や表現の訂正を同様に可能性の高い候補をクリックにより表示、選択することができる。

5. 今後の展望

今後の展開として提案する英文作成ツールの実装とその実際の使用例から、以下の三つの方向性が考えられる。

5.1 ツールの機能追加

ユーザが例文検索時にキーボード入力する日本語の語句に対し、言い換えや語彙などを提示する辞典データベースを利用することが挙げられる。実際の使用時にユーザが望むような例文が結果として得られないことも多かったためである。他言語の文章の作成の際、意識を行うことが多い。その支援を行うものとして、言い換え表現の辞書データの提示が必要であると考えられる。また短文コミュニケーションに利用する場合に、“LOL(Laughing out loud/爆笑)”や“IDK(I don't know/知りません)”のような略語、“XD(笑い)”や“<3(ハートマーク)”のようなエモティコンすなわち日本語で言うところの顔文字などの会話調の文章で用いられる定型句を組みかけ文に挿入できると、より意図を伝えやすいのではないだろうか。Urban Dictionary[17]のようなオンラインの投稿型スラング辞書を用いるのがよいかもかもしれない。

5.2 携帯端末への移植

単語オブジェクトに対するマウス操作による移動・編集やレイアウト作業は、キーボードによる文字入力を極力減らした形となっている。このようなマウスによるクリック、ドラッグ&ドロップなどを用いた操作手法は、指で直接ディスプレイに触れるタッチスクリーンを入力装置として搭載しているスマートフォンやタブレットPCなどの携帯端末に応用できるのではないかと考えている。

5.3 読む側の支援

提案したツールによる英文作成は、文法理解・構成においてユーザの英語能力に大きく依存している。そのため、ユーザが検索に用いた日本語の語句、用いた例文と単語オブジェクトはまさにユーザの意図する内容と実際に作成された英文とを繋ぐ重要な要素であるといえる。読み手が完成された英文からその要素を参照可能とすることで、より理解を促す支援ができるのではないだろうか。そして、一般的に書き手の意図する文と完成された翻訳文の間を繋ぐ機械翻訳を、例文データベースと置き換えることで、書き手と読み手が相互に参照できる多言語文章作成・閲覧の環境が構築できるのではないかと考える。

6. おわりに

本稿では、ユーザが手軽にデータベースの例文を引用し、文章として構築することができる英文作成ツールを提案し実装した。作成する英文に近い例文・表現の検索のための日本語入力にはキーボードを用い、その例文の単語や熟語などの移動・編集、レイアウトにはマウス操作を用いるものとした。これによりユーザは、英語知識が少なくても辞書データベースからの自動的な検索結果に頼りつつ、引用と訂正を書き手自らが行うことができる。実装したツール

を使用したところ、例文の検索からマウス操作による単語オブジェクトの移動を容易に行うことができた。また例文を引用し作成した英文でも他人が理解し、簡単なコミュニケーションに利用できるという感触が得られた。今後はより短文コミュニケーションに適合した機能を実装し、評価実験を行なう。また新たな可能性として、携帯端末への移植や例文による読む側の理解の支援などを行なっていきたい。

参考文献

- [1] Twitter, <http://twitter.com>
- [2] Usage of content languages for websites, <http://w3techs.com/technologies/overview/content.language/all> (2012.11.10).
- [3] Tumblr, <http://www.tumblr.com/>
- [4] 五十嵐健夫：文の構造を明示的に指定・表示することによる異言語間コミュニケーション、第19回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2011) 論文集, pp.114-119 (2011).
- [5] 松原茂樹, 加藤芳秀, 江川誠二：英文作成支援ツールとしての用例文検索システム ESCORT, 情報管理, Vol.51, No.4, pp.251-259, (2008).
- [6] 武田明子, 古郡廷治：例文をもとにした英文書作成支援システム, 情報処理学会論文誌, Vol.35, No.1, pp.53-61, (1994).
- [7] 大鹿広憲, 佐藤学, 安藤進, 山名早人：Google を活用した英作文支援システムの構築, DEWS2005, 4B-i8, (2005).
- [8] 池田佳代, 沼田秀穂, 金子正勝, 町田和彦：言語学習初心者に優しい他言語入力支援システムの開発とその評価, 情報アクセスシンポジウム 2010, article 3, (2010).
- [9] 綾塚祐二, 那和一成：GhostTweet: 短い文の入力のための新たなパラダイム, 情報処理学会研究報告-ヒューマンコンピュータインタラクション, Vol.2012, No.149(6), pp.1-5, (2012).
- [10] 辰川肇, Nigel Ward：キーボードとペンを併用するノートエディタ, インタラクション 2002 論文集, pp.209-216, (2002).
- [11] プログラミン, <http://www.mext.go.jp/programin/>
- [12] M. Resnick, J. Maloney, A. Monroy-Hernandez, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, and Y. Kafai : Scratch: Programming for all, Communications of the ACM, Vol.52, No.11, pp.60-67, (2009).
- [13] Alice, <http://www.alice.org/>
- [14] 緒方淳, 後藤真孝：音声訂正：選択操作による効率的な誤り訂正が可能な音声入力インタフェース, 情報処理学会論文誌, Vol.48, No.1, pp.375-385, (2007).
- [15] 緒方淳, 後藤真孝, 江渡浩一郎：PodCastle : ポッドキャストをテキストで検索, 閲覧, 編集できるソーシャルノートシステム, 第14回インタラクティブシステムとソフトウェアに関するワークショップ(WISS2006) 論文集, pp.53-58, (2006).
- [16] Google 翻訳, <http://translate.google.com/>
- [17] Urban Dictionary, <http://www.urbandictionary.com/>