

初級日本語作文における自己訂正 —コンコーダンスプログラム使用の試み—

伊東 克洋*

1. はじめに

第二言語ライティング（以下 SLW）における教師からの非直接的フィードバックに関する研究はこれまで様々なものが行われてきた。特に非直接的フィードバックは学習者に” guided learning and problem-solving”¹の機会を与え、” independent self-editors”²のスキルを育むことができると言われている。しかし、十分な言語知識やストラテジーを持っていない初級学習者においては非直接的フィードバックでは自己訂正が難しいということも同時に指摘されている。そのため、初級レベルにおける非直接的フィードバックを使った作文の自己訂正に関しては何らかのサポートが必要であると言えるだろう。本稿は初級日本語作文における自己訂正のサポートとしてコンコーダンスプログラムを使用する試みを報告するものである。コンコーダンスプログラムを使った言語学習はすでにくつかの研究でその効果が報告されている。しかしながら、学習言語を使ってプログラムに直接アクセスし、そのコーパスを分析するには、それに対応した言語運用能力が要求されるため、初級学習者におけるプログラムの使用に関する研究はあまり行われていない。本研究では、学習者にプログラムの使用感などを報告してもらうことで、初級レベルでのプログラム使用に対する可能性を模索し、今後の課題にしていきたいと考えている。

2. SLW におけるフィードバック

SLW におけるフィードバックは大きく分けて直接的フィードバックと非直接的フィードバックの二種類があると言われている。直接的フィードバックとは「教師が学習者の誤用に対し、正しい言語形式を明示的に提示すること」である³。直接的フィードバックにおいて学習者は教師に提示された正しい言語形式を確認またはコピーすることだけを要求される。一方、非直接的フィードバックとは「教師が下線、丸、コードやマークなど何らかの方法で誤用の存在を指し示すが、直接的に正しい言語形式は提示しない」ことであり、それらのヒントを基に学習者自身が誤用を訂正することを求められる³。この二種類のフィードバックの効果に関しては様々な研究で議論がなされているが、いくつかの研究ではより多くの認知プロセスの必要性があることから長期的な観点では非直接的フィードバックの方が習得につながりやすいと言われている⁴。しかしながら、自己訂正に十分な言語知識を有していない初級学習者に関しては効果が十分でない可能性も同時に示唆されている⁵。また、非直接的フィードバックに関しては誤用の指示に対する明示性の違いによる習得の差異についても議論されてきたが、十分な研究がなされているとはいいがたいだろう⁶。

3. データ・ドリブン・ラーニング

本研究で自己訂正のサポートとしてコンコーダンスプログラムを使用するに至った背景としてデータ・ドリブン・ラーニング（以下 DDL）理論が

* 米国パデュー大学大学院 博士課程

挙げられる。DDLとは第二言語学習者が対象言語のコーパスに直接アクセスし、それらを分析することでその機能や使用法などを帰納的に学習していくものである⁷。学習者はDDLにより認知プロセスを高めることができ、また経験を重ねることによって自律学習につなげていくことができる。しかしながら、学習者が何もコントロールされていないようなコーパスに直接アクセスし、分析できるような言語知識を持っていなければ、その効果は十分なものとは言えないだろう。

4. 初級学習者の作文における自己訂正

初級学習者の作文に対する自己訂正に関しては上述のような非直接的フィードバック、そしてDDL理論を組み合わせることによって十分な認知プロセスを経ることができ、効果的な習得につながるのではないかと考えられる。しかしながら、どちらの場合も相応の言語知識を要求されるため初級学習者にとっては難しいであろう。したがって初級作文に関してこれらを実践するためには、言語知識の不足をフィードバックにより補いつつ、コントロールされたコーパスに直接アクセスすることで帰納的な学習をしてもらう必要があると考えられる。

以上の理由から、本研究では初級学習者の作文に関して1)教師からの非直接的フィードバック、2)初級学習者用にカスタマイズされたコンコードプログラム⁸の2つの方法を用い、学習者に自己訂正を試みてもらった。

5. 調査対象

実験協力者は米国中西部にある州立大学で日本語を学習している大学生である(2セクション合計で22名、母語は英語、韓国語、中国語、ロシア語、アラビア語、ヒンディー語)。学習者のレベルは初級後半(外国語としての日本語を一年半程度

学習したレベル)で使用教科書はNakama 2⁸である。それぞれのセクションの均一性を測るためにSPOTを実施し、統計的に有意差がないことを確認した。

6. 実践方法

6-1. 非直接的フィードバック

本研究では非直接的フィードバックにおける明示性の違いを見るために下線とフィードバックコードの組み合わせと下線のみ⁹の二種類を使用した。フィードバックコードは以下の4種類である。

1) Wrong Conjugation (Conj.)

例) 書いていました(書いて)
きれくて、やさしいです(きれいで)
おもしろいでした(おもしろかったです)
 病気な人(の)

2) Wrong Form (F)

例) 彼は来ますと思います(来る)
 日本に行った前に(行く)⁹
 銀行は休みなはずです(休みの)
好きようです(好きな)
 郵便局はあけています(あいて)

3) Wrong/Missing Particles (P)¹⁰

例) 家に出る(を)
 日本のから手紙(からの)
 東京へ行きました(に)

4) Wrong/missing Conjunction (Con)

例) たくさん勉強しました。それからテストでいい点がとれました。(だから)
 学校に行きました。へだれもいませんでした。(けれども等)

6-2. コンコードプログラム

使用されたコンコードプログラムは筆者と

米国パデュー大学学部生の Andrew Schubert 氏によって本研究のために開発されたものである（図 1 参照）。使用されたプログラミング言語は Python、形態素解析にはオープンソースの形態素解析エンジン Mecab¹¹を使用した。また、Web ベースのプログラムであるため、インターネットが繋がらさえすればどこからでもアクセスすることができる。コーパスとしては Nakama の 1¹²と 2、そして Genki の 1¹³と 2¹⁴の計 4 つの教科書に含まれるすべての文（アクティビティーなども含む）を用いた。また、インターフェイスは日本語、英語どちらにも対応できるようになっている。

基本的な使用方法は検索ボックスに調べたい語彙や助詞、文型などを入れるとコーパスの中でそれらを含む文が短いものから順番にリストアップされる。動詞の場合は辞書形もしくは丁寧形（ます形）を入力すると活用された形もリストアップすることができる（過去形や否定形など）。また特定の活用を入力すれば（過去形など）活用された形のみが含まれた文がリストアップされる。

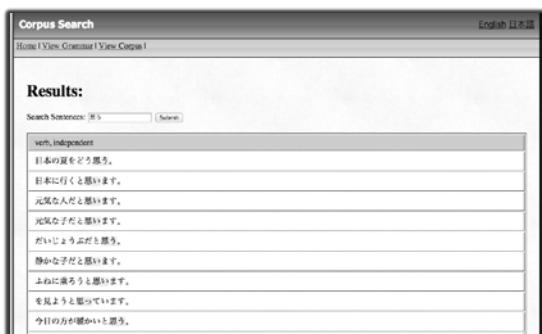


図 1. コンコーダンスプログラム

リストアップされたコーパスの上には形態素解析の結果（名詞、助詞、接続助詞、動詞、自立、一段など）が表示されるようになっている。また、コーパスに含まれる例文の不足などで検索が困難な場合を考え、View Grammar（文法項目から調べる）を使って、Nakama 1 と 2 の文法項目一覧から

対象となる文型を直接調べられるようになっている（図 2 参照）。

- Nakama2
 - Nakama2 Ch01
 - Describing what efforts are being made to attain a specific goal using ～ように
 - Expressing excessiveness using ～すぎる
 - Asking for and giving permission using ～てもいい
 - Expressing unacceptable actions or situations using ～てはいけない
 - Making a negative request using ～ないでください
 - Expressing capability using the potential forms of verbs
 - Giving suggestion using ～た方がいいです/～ない方がいいです
 - Giving suggestion using ～たらどうですか
 - Nakama2 Ch02
 - Using もう and まだ
 - Expressing conditions and sequence using たら
 - Expressing intention using the volitional form of the verb+と思う
 - Expressing intention using つもり
 - Expressing occasion and time using 時
 - Expressing intention using 予定
 - Nakama2 Ch03
 - Expressing the acceptability of an action or state using なくてもいい
 - Expressing chronological order using 後
 - Expressing chronological order using 前

図 2. 文法項目一覧

6-3. 実践手順

実験協力者が参加している日本語のクラスは一学期を通して Nakama 2 の 6 課から 10 課までを学習するため、各課の内容にあった作文のタスクを設定し、実施した。また、作文では新出語彙や文型などを積極的に使うように求められた。作文テーマは以下の通りである。

- 1) Ch.6 : 「私の一番思い出にのこる贈り物」
- 2) Ch.7 : 「私の好きな料理のレシピ」
- 3) Ch.8 : 「ついていない一日・出来事」
- 4) Ch.9 : 「異文化」
- 5) Ch.10 : 「気になること」

協力者はテーマに沿って、第一稿を 300 字以内で書き、教師（大学院生のティーチングアシスタント）¹⁵に提出した。教師は提出された作文をセクションごとに対象となる誤用にコーディングと下線、または下線のみで非直接的フィードバックを書き入れた¹⁶。添削された作文はクラス時間を使った 25 分程度のフィードバックセッション中に返却され、協力者はそれらを基に時間内に出来る限り自己訂正をするように求められた。ただし、検索方法などがわからない場合などは教師へ

の質問も許可された。なお、セッションは大学内のコンピュータラボを使って行われた。また、セッション中はコンピュータ上の不具合がないようにサーバ上からコンピュータの動きをチェックした。セッション終了後、作文を再提出し、教師は提出された第二稿に関しては直接的フィードバックを行った。最終的には各課のチャプターテストの際にテイクホームテストという形式で最終稿を提出させた。ただし、最終稿に関してはテスト形式のため、コンコーダンスプログラムは使用しなかった。

一学期を通してこれらの活動を行った後、学習者にはコンコーダンスプログラムを使った作文推敲に関するアンケートに答えてもらった。

7. 調査結果と考察

7-1. アンケート

紙面の都合上、アンケートで使われたすべての質問文について詳しく述べることはできないが、内容としては以下の通りである。

- 1) 従来の方法との比較
- 2) どのような点で効果的だったか
- 3) どのような点が難しかったか
- 4) 使用感はどうだったか

また、それらを測る指標として6段階のリッカート尺度を使った。使用した尺度の段階は以下の通りである。

- 尺度6：強く同意する
 尺度5：同意する
 尺度4：ある程度同意する
 尺度3：あまり同意しない
 尺度2：同意しない
 尺度1：全く同意しない

また、最後には良かった点と良くなかった点につ

いてコメントが書けるスペースを設けた。

7-2. アンケート結果

7-2-1. 従来の方法との比較

このパートでは従来の伝統的な紙媒体の辞書、教科書、そして協力者が普段使っているオンライン辞書などのオンラインリソースの三点について比較してもらった(表1参照)。表の中のパーセンテージはリッカート尺度の4から6までの合計である(紙面の都合上各尺度におけるパーセンテージは省略する)。紙媒体の辞書と教科書に関してはコーディンググループ(以下 Coded)、下線のみグループ(以下 Uncoded)ともにコンコーダンスプログラムの優位性を示している。しかしながら、オンラインリソースに関してはその優位性が下がり、特に Uncoded グループでは半数以上が従来のオンラインリソースの方が役に立ったと答えている。理由としては1)母語での訳や注釈などがついていること、2)普段使用しているため使い方に慣れていること、3)複数のサイトを同時に見ることができることなどが考えられる。また、学生に直接コーパスを分析し、認知プロセスを経て帰納的に学習してもらう目的のコンコーダンスプログラムでは、即座に結果を求める協力者のニーズには答えにくいことも事実であろう。プログラム自体の使用感の改善はもちろんだが、今後は教師側からの教育的意図を持った「手間」をどのように学習者に理解してもらうかという点も考えていかななくてはいけないと思われる。

	Coded	Uncoded
紙媒体の辞書	83.3%	90.0%
教科書	75.0%	90.0%
オンラインリソース	58.3%	40.0%

表1.従来の方法との比較(尺度6から4まで)

7-2-2. 効果的だった点

ここではコンコーダンスプログラムを使って推敲していく際にどの誤用に関して効果的だったかを答えてもらった(表2参照)。まず、語彙に関してはどちらもそれほど効果的ではなかったようである。これはコーパス自体が基本的に既習のものを前提に作られているので、新しく語彙を習得する機会は少なく、当然の結果といえよう。

一方で助詞、活用に関してはどちらも高い割合で効果的であったと評価されている。また、ここでは Coded、Uncoded グループで大きな差は見られなかった。本稿では実際にどの程度習得が進んだかという点に関しては触れていないが、学習者自身の感覚として、プログラムを使うことで(訂正したものが正しかったかどうかは別として)効果が得られたと感じている点は重要であろう。文法機能の帰納的学習がコンコーダンスプログラムの目的の一つであると考え、このアンケート結果はプログラムの有用性を示していると言えるかもしれない。

	Coded	Uncoded
語彙	50.0%	40.0%
助詞	83.3%	90.0%
活用	83.3%	80.0%

表2.効果的だった点(尺度6から4まで)

7-2-3. 難しかった点

ここでは特にコーパスとその表示方法について答えてもらった。まずコーパス中に使用された語彙については Coded、Uncoded グループともに難しさは感じなかったようである。これは上述したようにコーパス自体が既習語彙を中心に作成されているためであると考えられる。

次に調べたい語彙がコーパス上で見つからなかったと感じた協力者に関しては、Coded で 50%、Uncoded で 40%であった。データ上から検索ボックスに入力された語彙などを調べてみると多くの

漢語が検索されていることがわかった。これは特に中国語母語話者が中上級レベルの漢語を調べようとした結果であろう。今回のコーパスとしては初級教科書を使っているため、それらの漢語は含まれておらず、このような結果が出たと考えられる。

最後に検索後に表示される文の量に関してだが、これはどちらのグループ間でも意見がわかれており、ある個人にとって適切な量が他の個人にとっては多すぎる、または少なすぎるという可能性もあるため、母語・言語知識・レベルなどの個人的要因によってかなり左右される部分であると考えられる。

	Coded	Uncoded
語彙が難しい	25.0%	10.0%
語彙不足	50.0%	40.0%
表示される文(多)	50.0%	50.0%
表示される文(少)	41.7%	50.0%

表3.難しかった点(尺度6から4まで)

7-2-4. 使用感

ここではプログラムを使用した際の使用感について答えてもらった。まずはリストアップされた例文の中から自分が必要とする例文を検出する際の難しさに関してだが、Coded が 33.3%なのに対し、Uncoded は 60.0%の学習者が困難を感じたと答えた。これは非直接的フィードバックの明示性の違いによるものであると考えられる。例えば動詞の誤用に関して、Coded ではそれが活用なのか、形式の違いなのかが提示されるのに対し、Uncoded ではそういった指標がないために提示された例文に対しより慎重な分析が必要になるからであると考えられる。また、例文の検索の際に入力する語彙・言葉の選択も Uncoded の方が困難さを感じている理由も同様であろう。誤用の種類がコードを使って提示されていれば比較的検索する

語彙は絞りやすいが、下線のみであれば検索語彙を選択する際にすでにある程度の分析が必要になるからであると考えられる。しかしながら、実際に検出された例文が実際に推敲に役に立ったかという点に関しては、どちらのグループでも高い割合でその効果を感じていたようである。また、コンコーダンスプログラム自体の使用感についても Coded が 75%、Uncoded が 80% とともに高い割合を示した。これはインターフェイスを簡素化し、操作方法も可能な限りシンプルにしたことがその一因になっていると考えられる。

	Coded	Uncoded
対象文の検出	33.3%	60.0%
検索語彙の選択	41.7%	70.0%
検出された例文	75.0%	80.0%
全体的な使用感	75.0%	80.0%

表 4.使用感（尺度 6 から 4 まで）

7-3. コメント

アンケートの最後にこのコンコーダンスプログラムに対するコメントを書いてもらった。肯定的なものや否定的なものをそれぞれ以下にまとめる。

1) 肯定的な意見

- 正しい例が見られる
- 例文で正しい使い方を確認できる
- 一つの文法が複数の例文で見られる
- シンプルな例文がたくさん見られる
- 文法機能の確認ができるのがいい
- 例文が多いので同じ間違いをしなくなる
- 文型に合わせて、教科書の例文が一度に見られるのがいい

2) 否定的な意見

- 探している文がなかなか見つからない
- 例文が多すぎる／少なすぎる
- 期待したものと違う検索結果が出る
- コーパスに含まれている語彙がオンラインリソースより少ない
- 訂正すべきものがわかっている場合はいいが、あまり自信がないものに関してはなかなか見つからない
- 例文がシンプルなため、より複雑な文を作りたいときにはあまり役には立たない

肯定的な意見の傾向としては、プログラムを使うことによって、既習語彙を中心としたシンプルかつ正しい例文で文法機能が確認できるという点が挙げられる。これらは Google などの一般的なウェブベースの検索エンジンやオンライン辞書では実現されていない機能である。これは本研究のようなコントロールされたコーパスを使えば初級学習者でも簡単に適切なインプットを得つつ、帰納的な学習が可能になることを示唆していると言える。一方、否定的な意見として挙げられるのがコーパスの質と量であろう。これらは上述したような検索エンジンや辞書とは比べ物にはならないであろうし、個人的な要因によるものもあるため難しい点だと言える。帰納的な学習を目的にしているとはいえ、学習者が与えられた情報を適切に処理できなければプログラムを使う意味がなくなってしまうだろう。

8. 今後の課題

まず今後の課題としてあげられるのが検索方法のトレーニングの必要性である。本研究では一学期間、計 5 回のみプログラムを使ったフィードバックセッションを行った。その中で最初の 2 回はトレーニングセッションと位置づけ教師のアシトのもとでプログラムを使った推敲を試みた。しかしながら、それだけでは十分とは言えなかった

ようである。今後は十分なトレーニングを行った上で学習者が使用できる機会をもっと増やしていく必要があるだろう。また、そのためにも、今回のような推敲時だけではなく、第一稿を書く際などに使用することもプログラムの新たな可能性を測る良い機会になるのではないだろうか。

次にコーパス自体の問題が挙げられる。上述したようにコーパスの種類や質、量などは学習目的によって変わるものであるし、個人的な要因左右されるため、すべてに対応できるようなコーパスを作ることは難しいであろう。しかしながら、学習者の背景やニーズ、そしてレディネス等を事前に明確化し、可能な限りの最大公約数的コーパスを選ぶことが重要であると思われる。

今後はインターフェイスなども含め、初級学習者にとってより使いやすく、そして信頼性の高いプログラムに改善していけたらと考えている。また、他のオンラインリソースとの差別化ばかりを意識するのではなく、それらと共存・共有することで学習者にとってさらに便利で有用性の高いリソースになれるように研究を進めていくつもりである。

註

¹ Lalande, J.F., II (1982). Reducing composition errors: An experiment. *Modern Language Journal*, 66, 140–149.

² Bates, L., J. Lane, and E. Lange (1993) *Writing clearly: Responding to ESL compositions*. Boston; Heinle and Heinle

³ Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing* (pp. 81–104). Cambridge:

Cambridge University Press.

⁴ Frantzen, D. (1995). The effects of grammar supplementation on written accuracy in an intermediate Spanish content course. *Modern Language Journal*, 79, 329–344.

⁵ Ferris, D.R., & Hedgcock, J. S. (1998). *Teaching ESL composition: Purpose, process, & practice*. Mahwah, NJ: Lawrence Erlbaum.

⁶ Robb, T., Ross, S., & Shortreed, I. (1986). Salience of feedback on error and its effect on EFL writing quality. *TESOL Quarterly*, 20, 83–93.

⁷ Johns, Tim (1994) From printout to handout: Grammar and vocabulary teaching in the context of Data-driven Learning. In Odlin, Terence (ed.) *Perspectives in Pedagogical Grammar*. Cambridge; Cambridge University Press, 293–313

⁸ Hatasa, Y. A., Hatasa, K. & Makino, S. (2011). *Nakama 2: Japanese communication, culture, context* 2nd edition. Boston, MA: Houghton Mifflin Company.

⁹ 単純時制の誤用（昨日行きます等）は含まず。

¹⁰ 「は」と「が」の間違ひは含まず。

¹¹ <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

¹² Hatasa, Y. A., Hatasa, K. & Makino, S. (2010). *Nakama 1: Japanese communication, culture, context* 2nd edition. Boston, MA: Houghton Mifflin Company.

¹³ Banno E., Ikeda, Y., Ohno, Y., Shinagawa, C. & Tokashiki, K. (2011). *Genki I: An Integrated Course in Elementary Japanese* Second Edition. Tokyo: The Japan Times

¹⁴ Banno E., Ikeda, Y., Ohno, Y., Shinagawa, C. & Tokashiki, K. (2011). *Genki II: An Integrated Course in Elementary Japanese* Second Edition. Tokyo: The Japan Times

¹⁵ 2つのセクションでは別々の教師が教えていたが、inter-rater reliability では二人に差はなかった。また、教師は事前に筆者とともにフィードバックやコーディングなどに関するトレーニングをしてある。

¹⁶ 協力者に対象となる誤用の自己訂正に集中してもらうため、対象外の誤用に関しては直接的フィードバックを行った。