

英語学習用チャットにおける学習者コーパスと コンピュータ・ティーチング・アシスタント

Learner Corpus and the Computerized Teaching Assistant in Chat Sessions for Learning English

原田祐貨* 安田智宏**

*関東学院大学経済学部

**関東学院大学情報科学センター

Abstract: This paper describes the system of free online chat software called Language Education Chat System (LECS), which was developed for the benefit of English learners, teachers, and second language acquisition (SLA) researchers. Learners work in groups and simultaneously respond each other. Their teacher monitors the sessions and gives prompt feedback as well. All the written messages are stored in the database to build a learner corpus that offers linguistic samples to teachers who then can look for common errors worth mentioning in the class. It is also useful for researchers and textbook writers who can investigate familiar topics and typical ways of interaction among certain groups of learners. The system is based on SLA theories: learners may notice the gap between their interlanguage and English, test a hypothesis about comprehensibility or proper expressions, and internalize linguistic knowledge. To ensure this learning process, teachers can set a computerized teaching assistant to respond to errors, questions, and answers that were often made in the previous chat sessions. Useful research tools of word frequency and collocation lists combined with user-friendly layout make LECS popular software in Japan and abroad.

Keywords: chat, learner corpus, word frequency, collocation, error analysis, discourse analysis

1. LECS開発の目的

教育現場でのコンピュータの普及とメディア・リテラシーの向上とともに、英語学習にチャットを利用する教員も増えた。教室でグループ・ディスカッションを行ったときと比べて、チャットの方が参加者全員の意見を述べる機会が均等になった^[1]とか、学習者が互いの誤りを訂正しあう^{[2][3]}といった利点が報告されている。

しかし、ネット上に公開されている非学習用のチャットを教室で用いるにはいくつか難点がある。学習を妨害する侵入者を防げない、円滑な対話をできるように大人数の語学クラスを分割できない、学習者が入力した英文がデータとして残らないので復習や採点ができ

ない、などである。学習者と教員のためにこのような不便を解消することが、Language Education Chat System (以下LECS)の開発の第1の目的だった。

同じ英語学習者でも母語や文化、学習目的などの違いによって使う英語も異なる。そこで、学習者が書いたり話したりした英語を収集した膨大なデータ、すなわち学習者コーパスから見つけた学習者特有の誤りを改善するような教材が開発され始めた。そのためにはコーパスのデータは種類が多彩で量が豊富でなくてはならないが、一人の研究者が集められるデータには質量ともに限りがある。特に話し言葉のコーパスは録音した会話を書き取る手間がかかり、書き言葉に比べるとデータ量が極めて少ない。チャットは話し言葉ではないが、書き言葉よりも話し言葉に近く、データの偏りを補える。Web上に公開されてい

Taoka Harada* and Tomohiro Yasuda
Kanto Gakuin University
*E-mail:taoka@kanto-gakuin.ac.jp

るLECSは、全世界の教員がいつでも使え、研究者は公開の許可がとれたデータをどこからでも瞬時に閲覧し研究データとして使える。このように第二言語習得研究に有益なデータを収集し提供することが第2の目的である。

2. LECSの主な仕組みと機能

会話練習では各人の発言の機会を多くするために学習者を複数のグループに分けるが、LECSでも教員が同じ話題のチャット・ルームを複数用意し、各部屋を巡回して指導できる。チャットで入力された英文はデータベースに保存され、成績評価に活用できる(図1)。

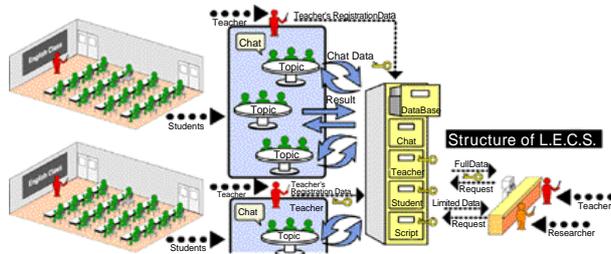


図1 LECSの概念図

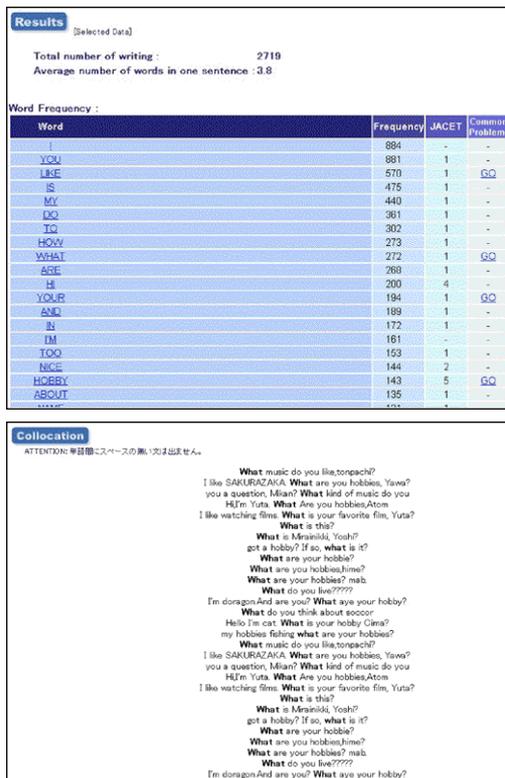


図2 使用単語の頻度(上)とコロケーション(下)

また、チャット終了直後に、単語の使用頻度表示(図2上)と各単語の使用例文をすべて表示したリスト(以下コロケーション、図2下)を見て、学習者が文の中で各単語をどのように用いたかを確認できる。単語の頻度表示とコロケーションを使うことで、より迅速に過去のチャットから教員や研究者が学習者に共通の誤りを見つけられる。またこれはオプション機能だが、誤りに対する教員の応答を予め登録すれば、チャットの最中にコンピュータ・ティーチング・アシスタント(以下CTA)が学習者の誤りを指摘する。研究者の便を考えて、過去のデータを話題別や学習者のタイプ別に検索できるようにした。

3. CTAのシステム構造

予め登録した単語やフレーズ、文(以下トリガー)を学習者が入力すると、それに呼応して応答文(以下レスポンス)をCTAが画面に表示する仕組みになっている(図3)。トリガーとレスポンスは1対1対応で、一つのトリガーに対し一つのレスポンスしか出せない。トリガーとレスポンスの組み合わせは、チャットを作成するとき教員が設定してもよいし、CTA Archivesにある組み合わせからも選べる。設定されたトリガーとレスポンスは、CTAのデータベースに保存される。

学習者がチャットに入力すると、コンピュータは最初に文字コードをチェックし、全角文字が含まれていると、CTAがWhat did you say?もしくは、What do you mean?と話しかけ、英語を使うように間接的に促す仕組みになっている。文字コードチェックの後、学習者の記述はCTAのデータベースと照合される。記述がトリガーと完全に一致する文字列を含むと、LECSはその記述と一緒にCTAのデータベースに記録されているレスポンス用の文章をチャットログに追加し、それを画面に表示

する。CTAのデータベースに登録すべきトリガーとレスポンスは、トリガーが多義語を含む場合は、レスポンスが文脈と合わなくなる危険があるので、登録には注意を要する。現在、CTAデータベースの参照によるサーバ負荷を軽減するため、一つのチャットにつき登録できるトリガーとレスポンスは最大20組までに制限している。

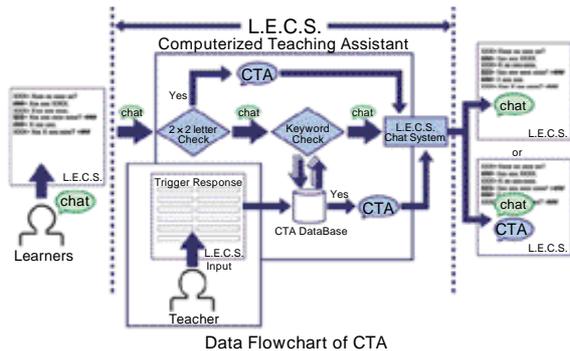


図3 CTAのシステム構造

4. CTAの設定方法

CTAのトリガーとレスポンスは、チャットを用意するときに設定する。過去のチャットで使用されたCTAの文例をまとめたCTA Archivesから文例を選んで使用できる。

誤りに対するCTAのレスポンスは、「誤り? + Do you mean 正しい表現」という構成で作ると、LECSはその後に学習者のチャット名前を自動的に付加して表示するので、学習者が確実に自分の間違いに気づいて正しい表現を学ぶことができる。ただしLECSには文脈を判断する機能がないので、レスポンスは、登録する前に必ず過去のチャットで使われた単語の連続に照らしていかなる場面でも不自然ではないかを確認する必要がある。

チャットの中で学習者がよく使った疑問文をトリガーにした場合は、「**疑問文の返答 + (具体例 +) 関連する疑問文**」という構造のレスポンスにすると、学習者が返答しやすい。またそのような疑問文に対する典型的な答えをトリガーにしたときは、「**(学習者の答**

え) + 詳しい説明を求める疑問文」を作れば、学習者からより具体的な答えを引き出せる。

5. LECSの指導方法と学習効果

LECSでは学習者が自分の英文を客観的に見直す機会を3回設けることで、正しい表現を意識するよう配慮した。最初はチャット入力欄に英文を入力したときで、次は送信後に他の参加者やCTAから間違いをチャットの最中に指摘されたときで、迅速な訂正は印象に残り正しい表現を覚えやすい。これらはLECS以外のチャットでも起こりうるが、より確実にするために、チャット終了後に自分の用いた英単語の頻度表示とコロケーションを見ることで、第3の見直しの機会を設けた。チャットの全文をダウンロードし、スペルチェック機能のついたワープロソフトにコピーすれば、学習者自身で綴りの間違いを直せる。

会話のクラスでは、教員が同時に複数のグループを見ることできないため、偶然通りかかったときにあった間違いは直せるが、どうしても断片的な指導に留まる。過去のLECSのデータから見つけた学習者特有の誤りをチャットの前に指導し、チャットの最中はその指導を教員が単独で（またはCTAと共に）行い、最後に単語の頻度表示とコロケーションを見て、指導を受けた言語表現を間違えていないかを確認することで、首尾一貫した指導が可能となる（図4）。

趣味に関するチャットでよく見られた文法上の誤りは、疑問文の語順、主語と動詞の形の不一致、可算名詞と冠詞の使い方、他の品詞を動詞化するなどである。チャットの最中は話の内容に関心が向き、学習者同士でこのような誤りを指摘し合うことは稀で、CTAや教員がはっきりと訂正する必要がある。例えばWhat do you like music?という語順の誤りに対して、You should ask, “ What MUSIC do



図4 LECSの利点

you like?” I like Irish music, like Enya. Have you ever listened to any of her songs?のように誤りの訂正と同時に話題を提供するとよい。

英語力が高くキーボード操作に慣れた学習者が話し相手を確認するために質問を繰り返し、英語力が低く不慣れな学習者がその質問に答えるのに終始するという傾向が見られた。返答は質問に対する直接的な答えに留まることが多く、趣味を問われても「音楽」や「スポーツ」とだけ答えて具体的な説明をしないので話題が発展しない^[4]。そこで、CTAが返答の遅い学習者に代わって英語力の高い学習者の質問に返答すれば、英語力の低い学生からの返答を待ち続けなくてよい。また、英語力の低い学生による具体的な情報に欠ける短い返答に対しては、CTAが詳しい説明を求めることで、彼らにより高度な表現に挑戦させることができる。その実例として、以下に好きなスポーツに関する男子大学生のチャットをあげる。プライバシー保護のために名前は偽名に変えたが、綴りなどの誤りはそのままにしてある。Tは教員を表し、数字はメッセージの順番を表す。

- 4 ICHIRO: Good! What is your hobby, JIRO?
- 5 CTA: I like playing baseball. I am a fan of the Yokohama Bay Stars. I wish Dai Majin were still playing in the team. Are you interested in baseball? >ICHIRO
- 9 JIRO: I like soccer. What your hobby? ICHIRO

- 10 ICHIRO: Yes, I like Bay Stars too. Who is your favorite player, CTA?
- 11 CTA: I play soccer myself and admire Hide Nakata. He is a genius! I think he is one of the best midfielders. >ICHIRO
- 13 T: ICHIRO, I like Yokohama Bay Stars, too. Do you think they will win the national championship this year?
- 14 JIRO: Hi, SABURO. I like soccer, too. What team do you like?
- 15 ICHIRO: Yes, of course. I believe.
- 16 ICHIRO: I like Juventus, because this team is always win, aren't you, JIRO?

5番目と11番目のCTAの返答は、What is your hobbyとWho is your favorite playerという疑問文をトリガーにしたレスポンスである。4番目のICHIROにJIROが返事をする(9番)までにCTAの他に二人がテニス(7番)とサッカー(8番)と答えたが、ICHIROはチームや選手に具体的に言及したCTAを選んで応答した(10番)。これは具体例が学生の応答を引き出しやすいことを示す。具体例を含んだレスポンスのもう一つの効用は、学習者が具体例に対する意見を述べるために、自分の言語知識を試したり他人の使った表現を借用することである^[5]。上記の引用例の後、学習者は好きなサッカー選手の話を始め、CTAにならって最上級を用いて批評した。最上級は高校で既習だがチャットの前に指導していないので、この場で既知の文法の実用例をCTAから転用したことを示唆している。

なお、CTAが不自然な応答をしたり対話に割り込む心配がある。11番目のCTAはサッカーの話をしており、ICHIROの野球の話と脈絡が合わない。このようにCTAは学習者の応答に必ずしも正しく返答できるとは限らないが、教員がCTAに代わって応答を修正できる

(13番)。複数の対話が入り乱れるチャットでは、やりとりの混乱が会話よりも多いが、対話が断絶することはない。教員がCTAの応答を修正したことで、ICHIROはCTAと教員の両方に答えなくてはならなくなったが、彼はまず教員に答え(15番)、それから返答を保留していたJIROに応答して話の流れをサッカーに戻している(16番)。

6. LECSの利用状況

前節で述べた学習上および研究上の利点が過去3年間で認知されるようになり、LECSの使い勝手の良さも報告されている^[6]。2003年5月末現在までのチャット数は1313室で、教員登録者総数は221名である。そのうち174名(79%)が英語を外国語として教えており(以下EFL)、第二言語として教えているのは(以下ESL)42名(19%)、ビジネス英語など特定の目的のために教えているのは(以下ESP)5名(2%)である。EFLは日本だけでなく中国、台湾、フランス、プエルトリコなど海外の語学学校や大学も含んでいる。またESLは主にアメリカの大学に併設の語学学校である。

学習者別に見ると、大学生が全学習者の63.8%を占め、短大生(12.1%)を含めると、全体のほぼ4分の3になる。また、1年生(大学生・短大生全体の87.7%)のうち、3分の2が必修英語である。高校生は全体の8.9%だが、学年別にその内訳を見ると2年生(76.2%)、1年生(20.2%)、3年生(3.4%)の順に多く、基礎的な語学力が備わった学習者に実際の英語を使う機会を与える目的でチャットを行っていると考えられる。

授業科目別に見ると、Speaking(29.4%)、Writing(20.3%)、Reading(19.3%)が全チャットの7割を占め、会話や作文指導の補助として使われていることがわかる。

7. 語学教育と研究への貢献

今後、教員同士が連絡を取り合って他大学の学習者と一緒にチャットをできるように、ユーザ・コミュニケーション用に教員談話室を設ける予定である。時差の少ない中国や台湾などの学習者と英語でコミュニケーションをとることは、教室の外で英語を使う機会の乏しい日本の教育現場では、国際語としての英語の言語運用力を試す有益な場となる。

一方、このシステムは英語教育用に開発されたが、文字コードを変えれば、ドイツ語、フランス語、ロシア語などの言語学習用チャットへの応用も可能で、言語教育の現場と研究に十分寄与できると考える。

参考文献および関連URL

- [1] Freiermuth, M. R.: Small Group Online Chat: The Great Equalizer. Teachers, Learners, and Computers: Exploring Relationships in CALL, JALTCALL, pp. 81-86, 1998.
- [2] Kitade, K.: L2 Learners' Discourse and SLA Theories in CMC: Collaborative Interaction in Internet Chat. Computer Assisted Language Learning, 13(2), pp. 143-166, 2000.
- [3] Warschauer, M.: Comparing Face-to-Face and Electronic Discussion in the Second Language Classroom. CALICO Journal, 13(2), pp. 7-26, 1996.
- [4] 安田智宏, 原田祐貨: 英語学習用のチャットにおけるコンピュータ・ティーチング・アシスタント. Local Decisions, Global Effects, JALTCALL, pp. 107-112, 2003.
- [5] Swain, M.: Three Functions of Output in Second Language Learning Principle and Practice in Applied Linguistics: Studies in Honor of H. G. Widdowson. Oxford University Press, pp. 125-144, 1995.
- [6] Freiermuth, M. R.: Internet Chat: Collaborating and Learning via E-conversations. TESOL Journal 11 (3), pp. 36-40, 2002.
- [7] <http://home.kanto-gakuin.ac.jp/taoka/lecs/>