

英語学習者の発話にみられる非流暢性に関する考察： 自己訂正と反復・フィラーの関係性

近 大志¹ 瀬戸口 彩花² 田中 悠介³ 神澤 克徳⁴

¹ 関西大学 ² 京都大学大学院 ³ 福岡大学 ⁴ 京都工芸繊維大学

{nanou7614, guchi.a.chan7, yusuke.tanaka.07, rnkp43470}@gmail.com

概要

第二言語学習者におけるスピーキング能力を測る指標の1つである流暢性を論じる際、それを阻害する非流暢性現象が注目されてきた。個別の非流暢性現象を習熟度と関連づける先行研究が多い中で、複数の現象間の関係性を大規模なデータから論じた研究は少ない。本研究では、英語学習者コーパスのKIT Speaking Test Corpus を利用し、4タイプの自己訂正（追加・削除・置換・倒置）に関して、反復・フィラーという他の非流暢性現象がどの程度共起しやすいかを調査した。その結果、追加・置換については習熟度ごとに共起傾向の違いが認められた。

1 はじめに

学習者のスピーキング能力を評価する一般的な指標として、複雑性 (complexity)、正確性 (accuracy)、流暢性 (fluency) の3つがある [1, 2, 3]。このうち、発話を淀みなく産出する能力を指す流暢性については、それを阻害する要因である非流暢性 (disfluency) の観点から分析が行われてきた [4]。

[4]によると、非流暢性を示す現象は、ためらい (hesitation) と修復 (repair) の2タイプに大別される。

ためらい (hesitation):

沈黙 (silent pause), フィラー (filler pause)

間延び (prolongation), 反復 (repetition)

修復 (repair):

自己訂正 (self-correction), 発話の中断 (false start)

第二言語学習の観点から非流暢性を扱った先行研究では、2つのタイプを区別した上で、特定の現象に着目しつつ、習熟度との関連性を分析する手法が多く採られている。例えば、英語学習者にみられる自己訂正のタイプを分類した [5, 6] や、その生起頻度と習熟度との相関を示した [7] が挙げられる。また、節の内外に生じたフィラーを習熟度の予測変

数として分析した [8] や、習熟度が高まるにつれて反復の生起頻度が減少したことを報告した [9] も挙げられる。

一方、学習者の発話を実際に確認すると、特定の非流暢性現象が単独で生起するのではなく、複数の現象が共起する例が多くみられる。[10]の自己修復 (self-repair) の心理モデルに照らすと、非流暢性現象としてのためらい・修復は、自己の発話中にある間違いに気づき、当該言語の規範に照らしつつ正しい伝達内容を計画し、それを実際に産出するまでの過程に伴う現象として捉えることができる。そのため、非流暢性現象を体系的に理解し、それを言語学習における指導法の提言へと活用するには、単独の非流暢性現象に特化するだけでなく、複数の非流暢性現象間の関わりをより大規模なデータから明らかにする必要が生じる。

本研究ではこのような課題に取り組むべく、ためらいのうち反復とフィラー、修復のうち自己訂正に着目し、両者の関連性を分析する。具体的には、次のリサーチ・クエスチョンを設定する。

RQ: 自己訂正と共起する反復・フィラーの生じ方に関して習熟度別の違いがあるか？

自己訂正には一定の認知負荷がかかる。一方、反復やフィラーが出現する動機づけの1つとして、話者への認知負荷が考えられる。両者の関係性を習熟度別に分析することで、学習者のレベルに応じてどのような自己訂正をする際に負荷がかかるのかを明らかにできると考える。

調査対象として、本研究では比較的規模が大きい英語学習者コーパスであるKIT Speaking Test Corpus (KISTEC)¹⁾のデータを利用した。そして、[5]や[7]で分析されている自己訂正のタイプ（追加・削除・置換・倒置）が発話中に生起するとき、自己訂正が完了するまでに反復・フィラーがどれだけ共起する

1) <https://kitstcorpus.jp/>

か、という観点からデータを分析した。

議論の流れは次の通りである。2節では調査に利用したコーパスデータの特徴およびデータの分析方法を明らかにする。3節では調査結果を示し、4節では結果に対する考察を与える。5節では結論と今後の課題を示す。

2 方法

2.1 分析データ

分析対象として、本研究では KISTEC を利用した。KISTEC とは、2018 年に京都工芸繊維大学の 1 年生が受験した CBT 方式の英語スピーキングテストの回答音声に、NICT JLE コーパスの仕様に準じた書き起こしとタグを与えたものである。本研究では、テキストに付与された 17 種のタグのうち、自己訂正タグ (<SC></SC>), フィラータグ (<F></F>), 反復タグ (<R></R>) に注目し、後述する手順からそれぞれの共起回数を調査した。

調査対象としたのは、自己訂正タグ (<SC></SC>) が付与された発話を習熟度の 3 グループから無作為に 1000 件ずつ抽出した 3000 発話である。習熟度グループの分類にあたっては、コーパスのデータに記載されているスピーキングテストのスコアを参照し、スコアと人数の分布に基づき、21-40 点のグループを初級者、41-50 点のグループを中級者、51-90 点のグループを上級者と設定した。

なお、自己訂正は 1 発話につき 1 つのみ生じる場合もあれば、複数生じる場合もある。以降の議論を簡潔にするために、自己訂正が複数生起する発話のうち、訂正後の表現が再び訂正の対象となったような発話については分析の対象外とした。

2.2 分析手順

本研究は [7] の基準に従い、自己訂正を 4 つのタイプに分類した。自己訂正前の表現を A、訂正後の表現を B とした場合、両者の関係は次のように定義される²⁾。

追加: A に 1 つ以上の形態素・単語を追加したものが B に相当する場合

削除: A から 1 つ以上の形態素・単語を削除したものが B に相当する場合

2) 置換は表現 A から何らかの要素を削除し、別の要素を追加したものと解釈ができる。そのため、[7] での扱いと同様に、削除・追加といった複数の操作を適用することで産出された表現 B に関しては置換と分類している。

置換: A の形態素・単語を 1 つ以上異なるものに置き換えたものが B に相当する場合

倒置: A の要素を変えず、語順を並べ替えたものが B に相当する場合

そして、それぞれのタイプにおいて、自己訂正が完了した部分までを範囲として、反復とフィラーの共起回数を調査した。例として、それぞれの自己訂正で反復・フィラーのいずれかが生じたものを (1) に挙げる。太字部分のうち、<SC></SC> に囲まれた表現は訂正前、それに後続するものは訂正後の表現に該当する。下線部分は、訂正が完了するまでに生じた反復・フィラー表現を指す。

- (1) a. 追加
... <SC>**a large place**</SC> <F>ehh</F>
<F>eh</F> **living a large place** ...
- b. 削除
<SC>**I was live in**</SC> <F>uh</F>
<R>I was</R> **I was in** Yamaguchi prefecture.
- c. 置換
... <SC>**she was**</SC> <R>she</R>
<R>she didn't</R> <F>uh</F> **she didn't** study hard.
- d. 倒置
Bill want to <SC>**more eat**</SC>
<F>uh</F> **eat more** and enjoy now.

3 結果

分類の際、もともとのタグ付けに問題があると筆者らが判断したものは分析から除外した。また、前節で述べたように、自己訂正が連続している例に関しても議論の都合上、分析から除外した。

まず、最終的に分析対象とした発話数と自己訂正の総数を習熟度で分けたものを表 1 に示す。

習熟度	発話数	自己訂正
初級	827	1098
中級	776	1172
上級	792	1143

次に、自己訂正が完了するまでに反復・フィラーが生じたかどうかを習熟度ごとに分類した結果を表 2 (初級者)、表 3 (中級者)、表 4 (上級者) にそれぞれ示す。

表2 初級者における自己訂正と反復・フィルター

	反復+フィルター	反復	フィルター	なし
追加	17	30	52	139
削除	3	7	15	25
倒置	0	0	2	0
置換	45	66	227	470
合計	65	103	296	634

表3 中級者における自己訂正と反復・フィルター

	反復+フィルター	反復	フィルター	なし
追加	6	7	22	54
削除	1	2	17	35
倒置	0	0	1	2
置換	37	57	256	675
合計	44	66	296	766

表4 上級者における自己訂正と反復・フィルター

	反復+フィルター	反復	フィルター	なし
追加	8	9	35	141
削除	4	3	18	45
倒置	0	0	1	1
置換	27	45	196	610
合計	39	57	250	797

表中の「反復+フィルター」は反復とフィルターがどちらも1回以上生起しているもの、「反復」は反復が1回以上生起し、フィルターが生起していないもの、「フィルター」は反復が生起しておらず、フィルターが1回以上生起したものの、「なし」はどちらも生起していないものを指す。

自己訂正の各タイプのうち、倒置は習熟度の違いや反復・フィルターの有無に関わらず、生起回数が非常に少ないことが判明した。この結果は [6, 7] と整合するものであり、以降の議論では、追加・削除・置換に限定した考察を行う。

自己訂正の各タイプについて、習熟度と反復・フィルターの共起のしやすさに関係性があるかどうかを確認するため、追加と置換に対してはカイ二乗検定を行い、データ数が5以下のセルを全体の20%以上含む削除についてはフィッシャーの正確確率検定を適用した。

その結果、追加と置換では有意差が認められた一方(追加: $\chi^2 = 14.631, df = 6, p = .023$; 置換: $\chi^2 = 28.792, df = 6, p < .01$)、削除では認められなかった ($p = .279$)。

次に、追加と置換に対する調整済み残差を表5と表6に示す。

表5 追加における反復・フィルターの調整済み残差

	反復+フィルター	反復	フィルター	なし
初級者	1.045	2.773**	0.457	-2.547*
中級者	0.341	-0.358	0.957	-0.769
上級者	-1.344	-2.581**	-1.217	3.226**

* $p < .05$, ** $p < .01$

表6 置換における反復・フィルターの調整済み残差

	反復+フィルター	反復	フィルター	なし
初級者	2.675**	2.774**	2.387*	-4.664**
中級者	-0.849	-1.071	-0.066	0.040
上級者	-1.734	-1.602	-2.264	3.575**

* $p < .05$, ** $p < .01$

4 考察

3節で示した結果より、習熟度からみた自己訂正と反復・フィルターとの関わりについては次のことが示唆される。

追加と反復・フィルターの関係:

自己訂正によって要素を追加する際、初級者は反復を伴う傾向にあり、上級者は反復・フィルターを伴わない傾向にある

置換と反復・フィルターの関係:

自己訂正によって要素を置換する際、初級者は反復・フィルターのいずれかを伴う傾向にあり、上級者はどちらも伴わない傾向にある

削除と反復・フィルターの関係:

自己訂正によって要素を削除する際、フィルターと共起する回数が他の組み合わせよりもわずかに高いものの、どの習熟度でも全体的な傾向は変わらない

ここからは自己訂正のタイプ別に考察を行う。まず、追加に関してである。[7]では初級者による追加の頻度が上級者よりも高い理由の1つとして、要素を漸進的に継ぎ足しながら伝達内容を構築するという方略が学習の初期段階で好まれるという可能性を挙げている。

先の主張を踏まえると、初級者の発話では、不完全な発話断片を徐々に完成させていく過程で、直前の表現を繰り返しながら追加要素の探索や確認を行っている可能性がある。例えば、初級者の発話である(2)では、最初の never に対して I can't を追加し、それが納得できるものであることを確認した上で、さらにそれに対して後続の発話を続けている。

- (2) a. So <SC>never</SC> <R>I can't never</R>
I can't never see ...

また、追加が生じた総数に関しては初級者が最も多く、次に多いのが上級者であり、最も少ないのが中級者であった。反復・フィラーとの共起傾向における違いがあるとはいえ、この結果は、熟達度が上がるにつれて追加の頻度が減少するとした [7] の報告とは必ずしも整合しない。

追加の生起回数が中級者で一旦減り、再び上級者で増えた理由としては、例えば、習熟度に応じて訂正内容に変化が生じる、といった可能性が考えられる。[11] は、日本人の英語学習者における自己訂正の方略を調査し、習熟度の低い話者は文法的な訂正を行う傾向にあり、上級者に近づくとつれ、伝達内容を意識した語彙的な訂正の割合が増えることを報告している。今回の分析対象としたデータにおいては、本研究で初級者と定めたグループでは冠詞 (a, the) ・不変化詞 (e.g., to, in) の挿入や動詞の屈折 (e.g., -s) といった、単一要素を追加した例が多い一方で、上級者のグループでは、内容語 (e.g., disadvantage is → computer's disadvantage is) や複数要素を一度に追加した断片 (e.g., talk to about → talk to my friend about) が比較的多く見られた。

次に、置換について考察する。分析結果からは、初級者による置換は反復・フィラーのいずれかと共起する傾向にあるが、上級者ではどちらも伴わずに要素を置き換える傾向にあることが分かった。今回の分類作業では [7] の同定基準を採用し、削除や追加といった複数の操作が関与しているような自己訂正を置換と認定している。このような定義の仕方には課題が残るものの、置換は自己訂正の中で比較的認知負荷の高い方略であり、学習初期段階にあたる話者は対象の言語規範を意識的に確認するために、置換の遂行時に他の非流暢性現象が伴いやすくなると考えられる。

また、助動詞や人称代名詞の置き換えといった文法的な側面に関わる置換 (e.g., don't → doesn't, we → our) が起こる際、上級者では反復やフィラーが伴うデータが比較的少数であった。その一方で、伝達内容の違いに大きく関わる内容語や動詞句といった置換においては、習熟度に関わらず、(3) のように反復やフィラーが共起する事例が多く観察される。

- (3) a. <SC>I wonder</SC> <F>uhh</F>
<F>uhh</F><R>I think</R> <F>uhh</F>

I think parents ...

- b. And <SC>these things are not</SC>
<F>uh</F> you cannot see these things ...

そのため、習熟度との関連から置換を詳しく論じる際、置換が文法的な誤りを契機とするのか、伝達内容の変更を動機としたものなのか、また、置換の対象が複数の語にまたがるのかどうかを区別した上で分析をする必要があると思われる。

削除に関しては習熟度ごとの差が小さく、反復・フィラーのどちらも伴わない割合がいずれも高かった。生起数が少なかったため参考程度ではあるが、新たな要素を考える上記 2 つの方略とは異なり、既に産出した要素の削除のみが関与するため、話者の認知負荷が少ない訂正であることが、習熟度に基づく有意差が観察されなかった理由と考えられる。

5 おわりに

本研究では第二言語学習者における複数の非流暢性現象の関連性を明らかにする目的のもと、自己訂正と反復・フィラーの共起傾向が習熟度によって違いがあるかどうかを確かめた。まとめとして、リサーチ・クエスチョンに回答を与える。

RQ への回答:

自己訂正のうち、追加と置換については反復・フィラーとの共起傾向が習熟度によって異なるが、削除では違いを認めることができなかった。

今回の調査にあたっては自己訂正前と訂正後の形式的な関係のみからタイプを分類したが、4 節での考察を踏まえると、訂正が対象言語の文法的な規範を意識したことで生じたのか、伝達内容の変更を意図したことで生じたのかを区別する必要があると考えられる。

また、今回の分析では自己訂正が連続した (4) のような事例を検討の対象外としたが、訂正の連鎖が一定のパターンを示す可能性も十分に考えられる。

- (4) ... <SC>Pokémon get</SC>
<SC>Pokémon</SC><F>eh</F>
<R>you</R> you get Pokémon.

これらの点を考慮した分析については今後の研究課題としたい。

謝辞

本研究は、文部科学省科学研究費補助金「英語スピーキングテストの採点効率化に向けた Task Achievement の解明」（課題番号：22K00736）を受けたものです。

参考文献

- [1] Peter Skehan. **A cognitive approach to language learning**. Oxford University Press, 1998.
- [2] Rod Ellis. **Task-based language learning and teaching**. Oxford University Press, 2003.
- [3] Rod Ellis. **The study of second language acquisition**. Oxford University, 2nd edition, 2008.
- [4] Simon Williams. **Disfluency and Proficiency in Second Language Speech Production**. Springer Nature, 2023.
- [5] Ann K. Fathman. Repetition and correction as an indication of speech planning and execution processes among second language learners. In Hans W. Dechert and Manfred Raupach, editors, **Towards a Cross linguistic Assessment of Speech Production**, pp. 77–85. Peter D. Lang, Frankfurt, 1980.
- [6] Kazuhiko Katagiri. Repetition, filled-pause, pause, and correction as an indication of speech planning and execution processes among japanese efl learners. **The Bulletin of the Kanto-Koshin-Etsu English Language Education Society**, Vol. 13, pp. 101–112, 1999.
- [7] 神澤克徳, 田中悠介, 井上優大. 第二言語の発話における自己訂正と学習者の習熟度の関連性. 言語処理学会第 25 回年次大会発表論文集, pp. 982–985, 2019.
- [8] Craig Lambert, Judit Kormos, and Danny Minn. Task repetition and second language speech processing. **Studies in Second Language Acquisition**, Vol. 39, No. 1, pp. 167–196, 2017.
- [9] Andrea Révész, Monika Ekiert, and Eivind Nessa Torgersen. The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. **Applied linguistics**, Vol. 37, No. 6, pp. 828–848, 2016.
- [10] Willem J.M. Levelt. Monitoring and self-repair in speech. **Cognition**, Vol. 14, No. 1, pp. 41–104, 1983.

- [11] 岩中貴裕. 日本人英語学習者の発話修正能力の発達. 神戸女子短期大学論攷, Vol. 53, pp. 9–28, 2008.