

英語シャドーイング音声評価データの分析

坪田 康[†] 伊藤 佳世子[‡]

[†] 京都工芸繊維大学 〒606-8585 京都市左京区松ヶ崎橋上町

[‡] 京都大学 〒606-8501 京都市左京区吉田本町

E-mail: [†] tsubota-yasushi@kit.ac.jp, [‡] cybel.kayoko@orion.ocn.ne.jp

あらまし 近年、シャドーイングが英語力の伸長に効果があるという研究報告が増えてきているが、人手での評価は非常に手間がかかるため、授業等で実施するのには容易ではない。シャドーイングの自動評価技術が簡単に利用できるになれば、効果的なシャドーイング活動の実施が容易になるであろう。自動化の予備的な検討として、A大学の1年生向けの英語授業（2クラス、63名）で、学習者の音声データを収集し、英語教員に留意すべき発音の確認、課題文中で発話できていない箇所の確認を依頼した。その評価結果に対して、課題文の難易度や学生が発話できない箇所の分析を行った。また、アライメント処理により、学習者音声モデル音声からどれくらい遅れて発話されているかの分析も行った。課題文の後半になればなるほど、遅れが増大する傾向がみられた。今後は、分析結果を踏まえ、アライメント処理の精緻化や課題文の再検討を行い、より詳細な分析を行う予定である。

キーワード 英語、シャドーイング、自動評価、教員による評価

An analysis on shadowing data evaluation of English

Yasushi TSUBOTA[†] Kayoko ITO[‡]

[†] Kyoto Institute of Technology Hashikamicho, Matsugasaki, Sakyo-ku, Kyoto, 606-8585 Japan

[‡] Kyoto University Yoshidahonmachi, Sakyo-ku, Kyoto, 606-8501 Japan

E-mail: [†] tsubota-yasushi@kit.ac.jp, [‡] cybel.kayoko@orion.ocn.ne.jp

Abstract Recently, many shadowing research proved that shadowing has the positive effect on improving English proficiency. In spite of the positive effect, it's not easy to introduce shadowing practices into classroom, since evaluations by teachers are labor and time consuming works. With the development of automatic shadowing evaluation system, the burden of those evaluation work would dramatically decrease and it would be easier to introduce more effective shadowing practices. As a preliminary experiment, we collected 63 students' shadowing speech in two classes for first year students at University A. Then, we asked three teachers to evaluate the following two things: 1. whether the important pronunciation is correctly pronounced or not, 2. mark the words which are not uttered at all. For those evaluation data, we set up two hypotheses: 1. At the timing of difficult words' appearance, the words, which are not uttered appear consecutively, and 2. as time passes by, words, which are not uttered decrease. Unexpectedly, we didn't find any supportive tendency for these two hypotheses. We also tried forced alignment with a speech recognition engine, called JULIUS and evaluated the delay between the model speech and students' speech, then we found that as time passes by, the delay increase. This implies that we can modify the second hypothesis. We will further investigate these hypotheses in the future.

Keywords English, Shadowing practice, Automatic evaluation, Evaluation by teachers

1. はじめに

聞こえてくる英語のモデル音声を、できるだけ即時的に口頭再生するシャドーイング練習は、ほぼ同時に

繰り返すため、英語特有の発音やリズム・イントネーションの習得が促進される。聞いた内容を理解しながら口頭発話していくので、リスニング技能とスピーキ

ング技能を高め、実践的なコミュニケーション能力を伸ばす効果があるといわれている。その効果の背景には、シャドーイングトレーニングにより、調音能力と音声知覚能力の向上、及び、音韻的ワーキングメモリ (phonological working memory)における内的リハーサルの高速化・効率化があることを示す研究成果もある。[1][2]。

シャドーイングの出来を点数化したり、誤りを指摘したりする評価・フィードバックが必要であり、英語レベルが適切で内容的に興味関心をもてる教材を多く提供することが重要である。シャドーイング練習のやる気を引き出すためには、どのように評価し、どんな教材を提供する必要があるかに関する体系的な調査や研究はほとんど行われていない。そこで、本研究では、英語のシャドーイング練習に対する学習意欲が高まり、自律的な学習が継続できることを目指して、(1) 学習者が客観的に自己評価できる自動評価システムを開発する。(2) 英語レベルが適切で興味・関心をもてる教材はどんな教材なのかを調査し、その調査結果に基づいて教材選択の基準を策定し、教材データベースを構築する予定である。

本報告では、その研究の最初の段階として、教員による評価に関する分析、個々の誤り部分に関する自動評価に関する検討について報告する。

2. シャドーイングアプリケーションの開発

今回のシャドーイング音声の収録に当たっては、共同研究者である東京大学峯松研究室にて開発されたアプリケーション[2]を利用した。どのような意識で開発されたアプリケーションかについて最初に説明する。

シャドーイング用アプリケーション：図1にシャドーイング用の Web アプリケーションのスクリーンショットを提示する。左上にある「シャドーイング開始」ボタンを押すと図の上にあるモデル音声の波形部分が再生されるので、学生はそれに続いて話をするという自動的に録音される仕組みである。なお、各文の最後には分の終了を示すために、「ピピピピ」という効果音が付与されている。また、同時再生、教師音のみ再生、録音のみ再生ボタンも付与されている。

収録用マイク：今回の収録は学習者の自宅を想定しているため、収録時の音質をできる限り向上させる必要がある。自宅の PC となるとサウンドシステムの精度にばらつきがあることが予想されるため、デスクトップ型の外付けマイクを貸し出すこととした。あわせて、収録用教示を web で提供することとした。マイクを PC に接続しても、PC 側で適切に選択しないとそれらを適

切に利用することは出来ない。USB マイクを使っているつもりが、実は PC 内蔵マイクで収録され、電源ノイズが乗ることは少なくない。また、息が直接かかる位置にマイクを置くとポップノイズが乗る。力んだ発声をする、波形振幅が AD 変換時の最大振幅を超えてしまうこともある。これらの現象を図示しながら説明している。

- 収録前の教示 <http://goo.gl/tXiKwA>



図 1：シャドーイング用アプリケーション

3. シャドーイングデータの収集・評価

3.1. シャドーイングデータの収集

A 大学の 1 年生向けの英語授業（2 クラス、63 名）で、学習者の音声データを収集した。2 で説明したアプリケーションを利用して収録させた。自宅に PC が無いなどの学生には、大学内にあるスピーキング練習用のブースで発話させた。

評価に使用しない 3 文で練習させた後、評価用の 15 文を発話させた¹。それぞれの文につき 3 回練習させた後、4 回目に発話した音声を評価用音声として採用した。ある程度モデル音声に慣れて、定常的に発話できるようになったものを評価対象とするためである。課題文一覧を付録 1 に示す。

3.2. シャドーイングデータの評価

3.1 の要領で収集した音声に対して、英語教員 2 名（日本人教員 1 名、英語母語話者 1 名）に、1. 課題文中で発話できていない箇所の確認、2. 留意すべき発音の確認を依頼した。評価に利用したシートを図 2 に示す。

課題文中で発話できていない箇所の確認として、図 2 の左側の部分を利用して、発音のよしあしはとまかく、

¹ 課題文は学習者のシャドーイング音声の誤り診断をするために、科研費 16H03447（自律的な英語シャドーイング学習を目指した自動評価と教材データベースの開発研究、代表：伊藤佳世子）で作成されたものを用いた。

発話ができている部分をマーカーで塗ってもらった。また、課題文を言えているかどうかに関して 5 段階で評価してもらった。図 2 の右側に留意すべき発音の一覧が並んでおり、それぞれの項目に関して、できているかどうかに関して判断をもらった。

1. に関して、課題文中の単語を横軸に、学習者の ID を縦軸に並べ、発話できなかった単語を赤く塗ったものが図 3 である。多くの学生が共通して発話できていない箇所があるのがわかる。そこで、下記の 2 つの仮説を立て、検証することとした。

- 仮説 1：難しい単語が出現すると、その語の直後から発話できない語が連続する
- 仮説 2：文の後半になればなるほど、発話できない語の出現率が向上する

評価者：	評価者 ID：			
No. 評価1 (発話できなかった部分を選択ペンで塗ってください。)	評価1 (★1)	評価2 (下線部、赤色部分に着目して評価してください)	チェックポイント	チェック箇所 評価2 (★2)
4 Your photo was really something.		Your <u>photo</u> was really something.	発音	photo something
5 The McDonald's house has been broken into.		The McDonald's house has been broken into.	発音	McDonald's
6 A policeman has come to check it out.		A policeman has come to <u>check it out</u> .	リズム	check it out
7 He said that they simply walked into the house.		He said <u>that they</u> simply walked into the house.	発音、連結	that they
8 The boy said that it had already been broken before he and his friend went to the house.		The boy said <u>that it had already been broken</u> before he and his friend went to the house.	リズム	that it has already been broken he and his friend
9 It was you who kicked the door open, wasn't it?		<u>It was you</u> who kicked the door open, wasn't it?	イントネーション	It was you wasn't it?
10 Did you just want to have a bit of fun, or were you trying to get some money?		Did you just want to have <u>a bit of fun</u> , or were you trying to get <u>some money</u> ?	イントネーション	a bit of fun some money?
11 The symptoms of fugu poisoning are a strange feeling around the mouth and throat and difficulty breathing.		The symptoms of fugu poisoning <u>are</u> a strange feeling around the mouth <u>and</u> throat and difficulty breathing.	発音	symptoms poisoning and breathing
12 Fugu is said to be so delicious that it has even started to be imported into Hong Kong and the United States.		Fugu is said to be so delicious <u>that</u> it has even started to be imported into Hong Kong and the United States.	発音	that be imported and
13 (Now then?) Tell me the truth.		(Now then?) Tell me the truth.	発音	truth
14 Shall we exchange some recent photos we've taken and discuss them on the Internet?		Shall we exchange some recent photos we've taken and discuss them on the Internet?		評価2のみ行ってください。
15 Because of the danger, fugu can only be prepared by chefs with a special license from the government.		Because of the danger, fugu can only be prepared by chefs with a special license from the government.		評価2のみ行ってください。
16 Most people who die from eating fugu these days are people who have tried their hand at preparing the fish themselves.		Most people <u>who die from eating fugu</u> these days are people who have tried their hand at preparing the fish themselves.	関係詞構文 読み文	who die from eating fugu these days
17 The prize money for each category is currently worth about a million dollars and the aim of the prize is to allow the winner to carry on working or researching without having to worry about raising money.		The prize money for each category is currently worth about a million dollars and the aim of the prize is to allow the winner to carry on working or researching <u>without</u> having to worry about raising money.	発音	without
18 Have you heard about the woman who put her wet dog in the microwave to dry and ended up cooking her dog by mistake? Or did you hear about the man who died at his desk at work, and nobody in the office noticed he was dead for five days? These stories have two things in common. They are both not true, and they are both urban legends.		Have you heard about the woman who put her wet dog in the microwave to dry and <u>ended up</u> cooking her dog by mistake? Or <u>did you hear about</u> the man who died at his desk at work, and nobody in the office noticed he was dead for five days? These stories have two <u>things</u> in common. They are both not true, and they are both urban legends.	音の連続	heard about put her end up did you hear about things in

図 2 学習者の発話チェックシート



図 3 学習者の発話チェック結果

4. シャドーイングデータの評価の分析

4.1 高難易度語と発話できない語の関係

仮説 1「難しい単語が出現すると、その語の直後から発話できない語が連続する」を検証するため、語集の難易度について調査した。測定には、JACET8000 をベースに単語の難易度を評する Word Level Checker(英文語集難易度解析プログラム)[4]を利用した。なお、評

価用の文には Fugu, McDonald などの固有名詞が含まれているので、これらは最も難しい語として処理することにした。

表 1 難易度の高い語

単語	難易度	単語	難易度
microwave	07	Photo	03
Chef	06	Poison	03
License	05	Delicious	03
Symptom	04	Import	03
Internet	04	Unite	03
Category	04	Urban	03
Currently	04	Legend	03
Spin	04		

固有名詞を除き、あまり難しい単語は使われていないことがわかる。評価結果から、発話できていない単語の上位 20 単語を取り出したものが下記のリストである。currently, legend が多少難しい単語として出てきているか、単独で難しい単語が影響を与えているわけではなさそうである。

表 2 発話できていない単語上位 20 単語

順位	%	単語	順位	%	単語
1	6.2	his	11	26.2	her
2	9.2	wet	12	32.3	he
3	15.4	at	12	32.3	legends
4	16.9	up	14	33.8	who
4	16.9	desk	15	36.9	said
6	18.5	her	15	36.9	said
6	18.5	ended	15	36.9	to
8	20.0	at	15	36.9	allow
9	23.1	urban	19	38.5	to
10	24.6	was	20	41.5	currently

難しい単語が現れて、その語が発話できないというよりは、後続する単語が発話できなくなっているのではと考え、2 つ組、3 つ組、4 つ組で発話できていないもので上位 15 単語を提示したのが表 3 である。

ここでも難易度の高い語が影響して、発話できなくなっていると思われる単語列は見つけられなかった。

表 3：発話できなかった単語組（上位 15 単語）

単語単位	割合	2単語組	数	3単語組	数	4単語組	数
his	6.2	at-his	51	at-his-desk	45	at-his-desk-at	39
wet	9.2	his-desk	51	his-desk-at	42	his-desk-at-work	31
at	15.4	her-wet	50	put-her-wet	35	who-put-her-wet	29
up	16.9	ended-up	50	desk-at-work	31	who-died-at-his	27
desk	16.9	desk-at	43	and-ended-up	30	about-the-woman-who	26
her	18.5	he-was	42	he-was-dead	30	woman-who-put-her	26
ended	18.5	urban-legends	39	who-put-her	29	died-at-his-desk	26
at	20.0	put-her	35	about-the-woman	28	her-wet-dog-in	25
urban	23.1	in-the	35	the-woman-who	28	about-the-man-who	25
was	24.6	at-work	34	her-wet-dog	28	wet-dog-in-the	24
her	26.2	said-to	33	died-at-his	28	hear-about-the-man	24
the	30.8	woman-who	33	hear-about-the	27	man-who-died-at	24
he	32.3	was-dead	33	the-man-who	27	the-woman-who-put	23
legends	32.3	and-ended	31	who-died-at	27	the-man-who-died	23
who	33.8	about-the	30	woman-who-put	26	are-both-urban-legends	23

そこですべての3つ組の単語列を時系列順にグラフ化(図4)を行った。難しい単語との関連で言えば、冒頭のFuguが多少の影響を与えている可能性はある。しかし、多くの課題文で冒頭以降に発話できない単語のパーセンテージは上昇しており、確定的なものとはいえない。仮説2の「文の後半になればなるほど、発話できない語の出現率が向上する」というのも、冒頭ではそのような傾向がみられるものの、課題文全体では必ずしもそうではない。

では、どういうときに発話できない語の出現率が増加傾向から減じるのかというと、一つはモデル音声のポーズが考えられる。多くの課題文は比較的長さが短いので、ポーズの影響はあまり見られないが、S17,S18等の長めの課題文では、モデル音声にポーズがあるところで、誤りの数が減じているのが観測された。息継ぎがあることで、意味処理の遅れなどがいったんリセットされ、新たにシャドーイングを開始できるようになっているのではと考えられる。

仮説1「難しい単語が出現すると、その語の直後から発話できない語が連続する」に関しては、少なくとも今回のデータからは顕著な例は見つからなかった。

仮説2「文の後半になればなるほど、発話できない語の出現率が向上する」に関しては、課題文冒頭ではそのような傾向はみられるものの、課題文全体では、上昇と下降を繰り返しており、仮説そのものを立証することはできなかった。一方で、モデル音声でポーズがあるところで、誤り率が減じている箇所があるためポーズを含め、他の要素を検討したモデリングは可能であると考えられる。

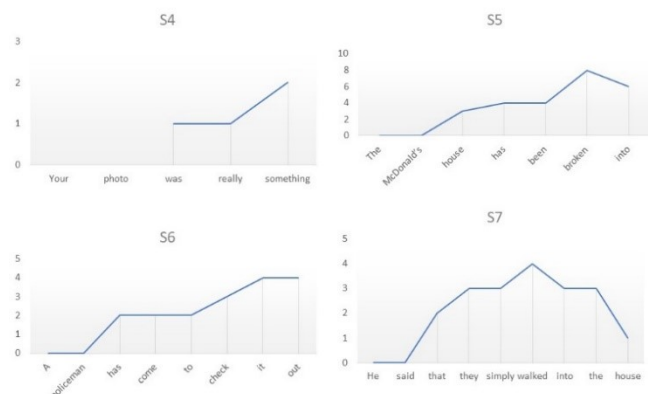
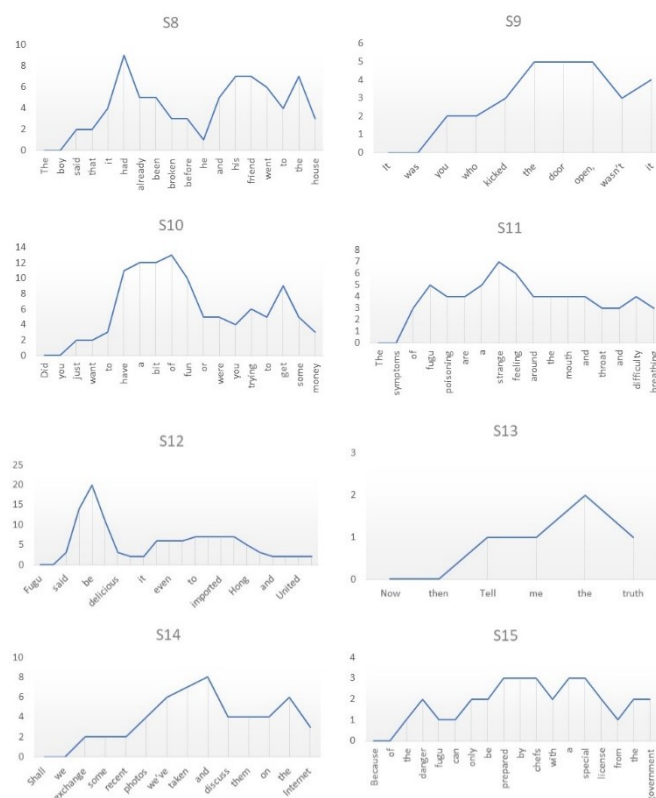


図4 課題文4-7の誤り率のグラフ



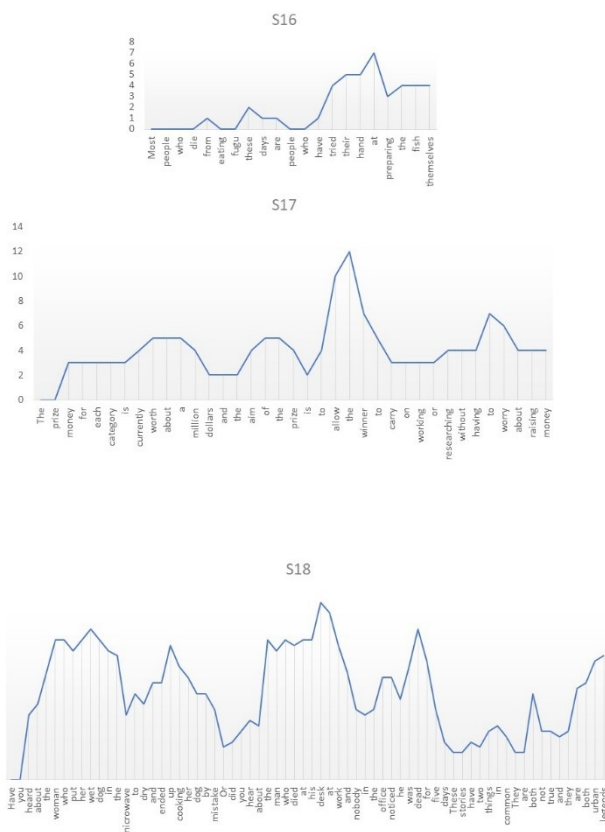


図5 課題文8-18の誤り率のグラフ

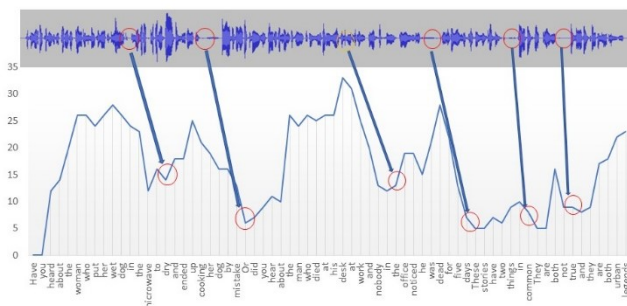


図6 課題文18の誤り率のグラフと
モデル音声のポーズ

4.2 発話率とモデル音声の長さ、速さの関係

課題文中で学生が発話できなかった単語数の課題文の全語数で割ったものを発話率と定義して、発話率とモデル音声の長さ、速さについての関係を分析した(図7)。発話率の順にソートしたのが図8である。図8の発話率で見ると、両端はS13とS18であり、S18は50%程度、S13は90を超えているが、後は75-80程度に収まっている。図7で見ると、S13のWPMが100以下で、S18のWPMは200を超えており、その他のものが100台であり、S8,13は特異な値であることがわかる。

	平均値	標準偏差	秒数	総語数	WPS	WPM
S4	82.5	14.2	1.9	5	2.6	157.9
S5	73.0	15.6	3.6	7	1.9	116.0
S6	83.8	7.2	3.0	8	2.7	161.1
S7	79.6	17.8	4.0	9	2.2	133.7
S8	75.0	13.3	6.4	18	2.8	167.7
S9	82.8	8.7	4.0	10	2.5	151.9
S10	75.0	14.1	5.7	18	3.1	188.2
S11	79.4	12.6	8.8	17	1.9	116.2
S12	75.4	17.3	9.2	22	2.4	143.3
S13	93.8	1.69	3.7	6	1.6	98.6
S14	76.6	15.3	4.8	14	2.9	176.1
S15	85.4	9.5	8.5	18	2.1	127.2
S16	82.4	13.7	9.1	21	2.3	138.2
S17	79.5	12.1	17.4	37	2.1	127.9
S18	50.6	19.7	19.6	68	3.5	208.5

図7 発話率とモデル音声の長さ、速さの関係

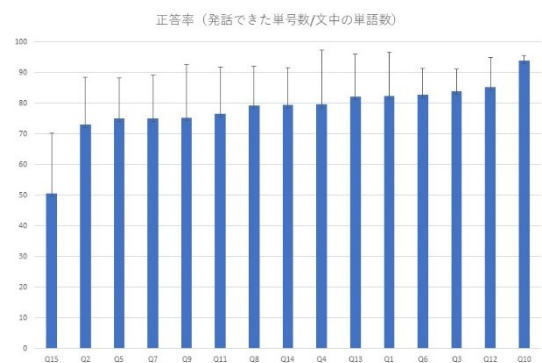


図8 発話率の分布

続いて、課題文毎に5段階評価したデータに対して、4,5の組(比較的良い評価)、3の組、1,2(比較的悪い評価)の分布を調べ、比較的よい評価のデータとWPM、秒数の相関を行った(図9)。S18はよい評価が少なく、S8はよい評価のものが多い。長さ、秒数、WPM、秒数の相関を行った(図9)。S18はよい評価が少なく、S8はよい評価のものが多い。長さ、秒数、WPM、秒数の相関を行った(図9)。

	4,5	3	1,2	WPM	秒数
S4	75.9	19.0	5.2	157.9	1.90
S5	55.2	29.3	15.5	116.0	3.62
S6	81.0	13.8	5.2	161.1	2.98
S7	67.2	22.4	10.3	133.7	4.04
S8	51.7	31.0	17.2	167.7	6.44
S9	74.1	17.2	8.6	151.9	3.95
S10	69.0	19.0	12.1	188.2	5.74
S11	62.1	27.6	10.3	116.2	8.78
S12	66.7	22.8	10.5	143.3	9.21
S13	93.1	3.4	3.4	98.6	3.65
S14	63.8	27.6	8.6	176.1	4.77
S15	82.5	12.3	5.3	127.2	8.49
S16	69.0	27.6	3.4	138.2	9.12
S17	54.4	35.1	10.5	127.9	17.36
S18	15.8	36.8	47.4	208.5	19.57

図9 課題文の5段階評価の分布と長さ、速さとの相関

4.3 学習者音声の遅れの分析

仮説 2 と関連して、文の後半になればなるほど、認知処理は増大するのではないかと考え、個々の学習者の発話がモデル音声の発話からどの程度遅れているか、単語単位で調査した (図 10)。縦軸に時間 (msec)、横軸には、課題文の単語を横に並べている。なお、アライメント分析には、音声認識エンジン JULIUS[5]を用い、モデル音声と学習者音声に対して Forced Alignment をとり単語開始時間の差分をとった。

まずは全体の結果を概観する。一部でマイナスの値があるものの、それを除けば、概ね、1 秒までの遅れでシャドーイングがなされている。これは多くのシャドーイングの参考書で概ね一秒以内にモデル音声を繰り返すようにと指導されていることと一致する。また、細かいところで増減はあるものの、課題文の後半になればなるほど遅れが増大している。この結果からだけでははっきりとしたことはいえないものの、認知処理が追い付かなくなっている傾向があることを示唆するものと考えている。短い時間で遅れが増減しているのは、音節数がばらばらの単語が並んでいる、機能語、内容語が混ざっているため、多少の揺れが生じているものと考えている。

続いて、マイナスの値をとっているものについて述べる。S12,S17 では最後の語で、S13 ではすべての単語で、マイナスの値が出ていた。学習者がモデル音声より先行して発話していることになり、シャドーイングの定義に反する。実際のデータを精査してみたところ、S12,S17 ではシャドーイング後にノイズがあることが多かったためか、Forced Alignment の処理でノイズ部分まで単語を発話しているとする結果が多かった。今後、アライメント処理の精緻化を検討したい。S13 は単語数が少なかったためか、実際にモデル音声よりも早く発話していた。シャドーイングの練習にならないので、モデル文の再選定を行うことで改善したいと考えている。

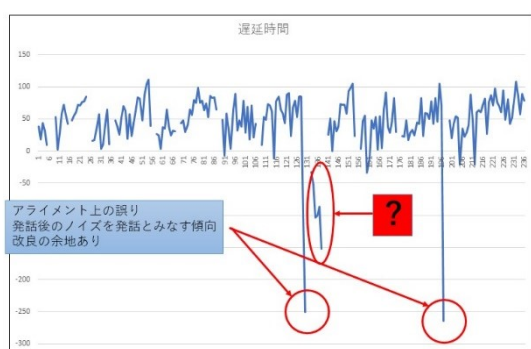


図 10 モデル音声からの時間遅れの分析

5. おわりに

英語授業 (2 クラス, 63 名) の学生に対し、専用のシャドーイングアプリケーションを利用して、学習者の音声データを収集し、英語教員による評価を行った。難しい単語の出現が契機となり発話できなくなる、文の後半になればなるほど発話できなくなると仮説を立てたが、今回の実験に関してはそのような傾向はみられなかった。モデル音声に息継ぎがあると誤り率が下がっていることはあった。WPM が 200 を超えると発話率が極端に下がったり、100 を切ると極端に上がったという現象は見られたものの、これらはどちらかというとも極端な例であると考えられる。

また、Forced Alignment により学習者発話がモデル音声からどれくらい遅れているかを分析したところ、文の後半になればなるほど、時間遅れが増大している傾向が見られた。今後はアライメント処理の精緻化、課題文の再選定を行ったり、学習者の傾向をよりとらえやすくなるような実験を行うなどにより、自動化のための検討を行っていくつもりである。

本研究は科研費 16H03447 (自律的な英語シャドーイング学習を目指した自動評価と教材データベースの開発研究、代表：伊藤佳世子) の助成を受けたものである。

参考文献

- [1] 門田修平, “シャドーイング・音読と英語習得の科学 インプットからアウトプットへ,” コスモピア, 2012.
- [2] 峯松信明, 楽俊偉, 山内豊, 伊藤佳世子, 齋藤大輔, “多人数同時発声環境における効果的なシャドーイング音声収録に関する検討,” 日本音響学会講演論文集 (秋), 3-Q-27, 2016.
- [3]
- [4] 染谷 泰正, “オンライン版「英文語彙難易度解析プログラム」(Word Level Checker)の概要およびその教育研究分野での応用可能性,” 青山学院大学紀要 51, 99-122, 2009
- [5] 李 晃伸, “大語彙連続音声認識エンジン Julius ver. 4,” 電子情報通信学会技術研究報告. SP, 音声 107(406), 307-312, 2007-12-13

付録 1：課題文一覧

- S4: Your photo was really something.
S5: The McDonald's house has been broken into.
S6: A policeman has come to check it out
S7: He said that they simply walked into the house.
S8: The boy said that it had already been broken before he and his friend went to the house.
S9: It was you who kicked the door open, wasn't it?

S10: Did you want to have a bit of fun, or were you trying to get some money?

S11: The symptoms of fugu poisoning are strange felling around the mouth and difficulty breathing.

S12: Fugu is said to be so delicious that it has even started to be imported into Hong Kong and the United States.

S13: Tell me the truth.

S14: Shall we exchange some recent photos we've taken and discuss them on the Internet?

S15: Because of the danger, fugu can only be prepared by chefs with a special license from the government.

S16: Most people who die from eating fugu these days are people who have tried their hand at preparing the fish themselves.

S17: The prize money for each category is currently worth about a million dollars and the aim of the prize is to allow the winner to carry on working or researching without having to worry about raising money.

S18: Have you heard about the woman who put her wet dog in the microwave to dry and ended up cooking her dog by mistake? Or did you hear about the man who died at his desk at work, and nobody in the office noticed he was dead for five days? These stories have two things in common. They are both not true, and they are both urban legends.