様式 C-19

科学研究費補助金研究成果報告書

平成２３年６月２４日現在

機関番号：３４３１５
研究種目：基盤研究（C）
研究期間：2008～2010
課題番号：２０５２０５３７
研究課題名（和文）　英語ライティング評価手法の開発

研究課題名（英文）　Text Profiling of English Writing

研究代表者
朝尾　幸次郎（ASAO KOJIRO）
立命館大学・文学部・教授
研究者番号：４０１０２４６２

研究成果の概要(和文)：英語を外国語として学んだ人で正確に文章を書くことのできる人でも、その英文の英語らしさ、上手下手には大きな違いがみられる。また、英語母語話者の書く文章でも、わかりやすさ、読みやすさには大きな違いがある。本研究では、このような違いが何に由来するものかを探った。このため、テキストの語彙、文長、語連続に着目し、テキストの英語らしさ、難易などの性質にかかわる要因を探り出し、それを一覧表として提示する手法を提案した。
研究成果の概要（英文）：Even advanced learners of English differ extensively in their proficiency of writing. Some write in near-native capacity while others leave a number of traces of “foreign accents” in their writing. Grammatical accuracy does not necessarily endorse native fluency. Native speakers of English also differ greatly in the readability of the text they write. This study investigated English texts in order to identify the sources of stylistic variability by focusing upon the vocabulary, sentence length and collocations of the text.

交付決定額

（金額単位：円）

	直接経費	間接経費	合　計
２００８年度	800,000	240,000	1,040,000
２００９年度	500,000	150,000	650,000
２０１０年度	500,000	150,000	650,000
総　計	1,800,000	540,000	2,340,000

研究分野：　応用言語学
科研費の分科・細目：　言語学・外国語教育
キーワード：　英語、ライティング、熟達度、プロファイル、コーパス

１．研究開始当初の背景

(1) 外国語を書く力の評価基準を作成することは米国、欧州でさかんに進められてきた。そのひとつが全米外国語教育協会(ACTFL)の基準である。これは《もっとも初心者》についての「アルファベットの文字を書くことができる」から始まり、《中級の中レベル》「短い、簡単な手紙を書くことができる」といった例による詳細な基準である。

しかし、これは現場では役に立たない。英語の作文を読んで、これがどのレベルに相当するか、印象的に判断はできる。しかし、どこに弱点があり、どこに不足があるのかを示すことはできない。

また、「手紙を書くことができ」ていても、英語として自然でない表現が多い手紙、短い文ばかりで書かれた手紙はどう判断すればよいのだろうか。実際には判断に迷う。

(2) 全米外国語教育協会(ACTFL)など海外のライティング到達指標に欠けているのは英

語としての正しさ、英語らしさの評価である。リーディングと異なり、ライティングでは学習者はみずから言語的正しさ、適格性を生み出さなければならない。

「安全だ」というつもりで書いた生徒の英語、*It is safety.「私の家は 4 人家族です」というつもりの *Our family is four. は文構造という面では正確であるものの、英語としては誤用であり、英語力、ライティング力の未熟を示すものだ。このような、私たちが一番判断の基準がほしいところが ACTFL の基準ではすっぽり抜け落ちている。これは基準を立ててきたのが英語母語話者であるという理由によるところが大きい。外国語を学び、教える者の立場からの新しい基準、視点が必要だ。

２．研究の目的

(1) 私たちが書く英語の上手、下手はどのように判定すればいのだろう。TOEFL のライティング試験では採点者の主観的判断により 1〜5 のレベルにおおまかに区分している。おおまかにしか判定できないのは主観テストという制約のほかに、《英語のライティング力を形成する要因とは何か》というその詳細があきらかになっていないという理由によるところが大きい。

そこで考え方を変えてみる。ライティング力を構成する要因を網羅し、それを調べあげることが現実的でないのであれば、学習者の書く英語は英語母語話者が書く英語と比べてどのような点で異なっているかを探ってみる。語彙の豊富さ、繰り返し、文の長さ、連語を基準に、英語母語話者の書く英語との違いを一覧表にして提示してみる。いわば、ライティング力プロファイルの提示である。英語ライティング力の詳細はそれを見て、私たちが判断すればよい。

(2) 英語ライティング力のプロファイルを提示するしくみの構築が目標ではあるものの、最終的にめざすものは「英語を書く力とは何か」「それを構成する要因は何か」をあきらかにすることである。英語母語話者の書く英語との対比のなかで、影響の大きい要因、そうでない要因がうかびあがってくる。そこから英語ライティング力を構成する要因をあらためて定義する。

３．研究の方法

(1) 私たちの書く英語は、文法的にはまちがってはいないけれども英語としては自然ではないものがある。英語母語話者が書く英語であっても、同じ内容を表現していながら、テキストとしてのやさしさ、むずかしさに違いがある。このようなテキストの性質を決定する要因を探るため、使用されている語彙の種類と量を比べる。

(2) テキストの難易を決定する指標のひとつとしてとりあげられることの多い、語数で数えた文長をテキストごとに比較し、検証する。

(3) 英語らしさがもっとも端的に表れるのは語連続の性質である。このため、2 語連続、3 語連続、4 語連続についてテキストを比較する。

４．研究成果

(1) テキストの難易を測る指標として知られている Flesch および Dale-Chall の公式の検証を行った。Flesch のリーダビリティー公式創案の根拠となったものは、文の平均単語数、接辞の数、人への言及の回数の 3 つであった。この公式による読みやすさの評価はおおむね人々の印象と合致していて評価を得た。しかし、印象とは一致しないものもあった。この公式によれば、*Reader's Digest* は *New Yorker* よりも読みやすいと判定される。ところが、教育を受けた人の間では *Readerc's Digest* は退屈で、*New Yorker* の方が 10 倍も読みやすいと考えた。

このような反省から Flesch は公式を見直すことを始めた。語数で数えた文長は読みやすさの点からきわめて重要なものであるものの、それが過大視されがちである。接辞は数え出すのが困難である点で不安定である。人への言及も恣意的であるという意見もあった。

これをもとに Flesch がリーダビリティー算出の基礎として使ったのは文の平均単語数、音節数で数えた単語の長さ、人に関する単語の割合、人に関する文の割合の 4 つである。最後の「人に関する文の割合」は物語文とそれ以外の文章を区別するための方法である。引用符で囲まれた文は人の発話で、これらは一般にやさしい表現であることが多い。このようにして算出されたのが Flesch のリーダビリティー公式である。

これに対し、Dale-Chall の考え方はやや異なっていた。Dale-Chall の公式は「語数で数えた文長」と「基本 3,000 語を定め、これ以外の語が含まれる割合」を基礎としている。

Dale-Chall の考え方は基本的に Flesch のものと同じである。Flesch がとりあげている「音節数で数えた単語の長さ」ということは長い単語、すなわち、ラテン系の語彙ということである。一般にラテン系の語彙は日常的に使用するゲルマン系の語彙と異なり、むずかしいとされる。ちょうど日本語における漢語と和語の違いに似たところがある。日本語においても和語中心の文章は日常的な表現であるのに対し、漢語が多い文章は学術的である。

英語における文章の読みやすさにかかわ

る要因は、突きつめれば文長と語彙になるようだ。次はルイス・キャロルの『不思議の国のアリス』の一節である。涙の池に落ちたアリスと生き物たちが陸に上がり、そこでドードーが知ったかぶりをして言う場面である。

‘In that case,' said the Dodo solemnly, rising to its feet, ‘I move that the meeting adjourn, for the immediate adoption of more energetic remedies—'
‘Speak English' said the Eaglet. ‘I don't know the meaning of half those long words, and, what's more, I don't believe you do either!' And the Eaglet bent down its head to hide a smile: some of the other birds tittered audibly.

ここでドードーのことばがむずかしいのは小鷲が long words と言っている、ラテン系、ギリシャ系の語を多用している点である。

(2) テキスト比較の対象として Howard Zinn よるアメリカ史のテキストを選んだ。これには若い世代向けに書き直した版がある。同一の内容のものを基本的にセンテンス単位で書き直している点で、本研究の目的に好個の素材である。以下、元の版を「原典」、やさしく書き直した版を「書き直し版」として示す。

(3) 書き直し版で広く使われている方法のひとつは、次のように原典の文をふたつに分け、短くするというものである。
（原典）Arawak men and women, naked, tawny, and full of wonder, emerged from their villages onto the island's beaches and swam out to get a closer look at the strange big boat.
（書き直し版）Arawak men and women came out of their villages on to the beaches. Full of wonder, they swam out to get a closer look at the strange big boat.
文をふたつに分けるというのは単に短くするだけではなく、スキーマを豊かにする意図もみられる。次の原典の例はただ and で文をつないだだけのものである。しかし、書き直し版を見ると、文をふたつに分けたのは and で表された背景知識にかかわる意味を明示的に示したものであることがわかる。
（原典）For, like other informed people of his time, he knew the world was round and he could sail west in order to get to the Far East.
（書き直し版）Like other informed people of his time, Columbus knew that the world was round. This meant that he could sail west from Europe to reach the East.

文をふたつに分けるということは単に文長を短くするだけではなく、スキーマを豊富化させるなど、内容にかかわる意図もあることに留意しなければならない。このため、次の例のように書き直し版では原典よりも逆に文長が長くなった例もある。
（原典）Thirty-three days after leaving waters known to Europeans, Columbus and his men saw branches floating in the water and flocks of birds in the air.
（書き直し版）It was early October 1492, and thirty-three days since he and his crew had left the Canary Islands, off the Atlantic coast of Africa. Now they saw branches and sticks floating in the water. They saw flocks of birds.
原典はスキーマを背景にし、緩急のあるドラマチックな表現であるのに対し、書き直し版は直接、事実を語る表現になっている。

(4) テキストの難易は形だけで判断できるものではない。原典と書き直し版では同じ内容を記述していても、情報が一方に欠落している個所も多くみられる。次は書き直し版で原典の情報が落ちている例である。
（原典）When Columbus and his sailors came ashore, carrying swords, speaking oddly, the Arawaks ran to greet them, brought them food, water, gifts.
（書き直し版）When Christopher Columbus and his soldiers came ashore, carrying swords, the Arawaks ran to greet them.
ここでは長い文をふたつに分けるのではなく、枝葉にわたる情報を落としていることに留意する必要がある。

(5) 書き直し版では代名詞を避け、名詞で表現する例が多い。
（原典）He later wrote of this in his log.
（書き直し版）Columbus later wrote about the Indians in his ship's log:
He を Columbus で、this を the Indians で具体的に表している。

(6) 原典よりも書き直し版が長くなった例では、おおむねスキーマを豊富化させる意図がある。文の長さとわかりやすさの関係は機械的なものではない。
（原典）The information that Columbus wanted most was: Where is the gold?
（書き直し版）As soon as he arrived in the islands, he seized some Arawaks by force so that he could get information from them. The information that Columbus wanted was this: Where is the gold?

(7) 平易な語彙を使う、難語を避けるという

のも広く使われる方法である。典型的な例はラテン系の語をゲルマン系の語で言い換える方法である。次では persuade を talk into で、finance を pay で表している。
（原典）He had persuaded the king and queen of Spain to finance an expedition to the lands, the wealth, he expected would be on the other side of the Atlantic--the Indies and Asia, gold and spices.
（書き直し版）Columbus had talked the king and queen of Spain into paying for his expedition.
　この例はきわめて多く、これまでのリーダビリティー公式が背景にしている考え方を裏書きするものである。

(8) 関係代名詞の埋め込みを避けることがテキストの平易化ではよくみられる。
（原典）He then sailed to what is now Cuba, then to Hispaniola (the island which today consists of Haiti and the Dominican Republic).
（書き直し版）He sailed to several other Caribbean islands, including Hispaniola, an island now divided between two countries, Haiti and the Dominican Republic.

(9) センテンスあたりの平均語数とその標準偏差は原典で、19.21、10.79、書き直し版で14.58、6.77である。これによってもFleschなどのリーダビリティー公式が背景にした考えがもっともなものであることが裏付けられる。

(10) 語彙頻度について特徴的なのは、書き直し版における人名、地名の多さである。これは書き直し版では代名詞を使わず、照応関係の煩雑さを避けたことによる。たとえば、Indians の出現回数は書き直し版では原典の倍近い。

(11) 代名詞についても頻度に違いがみられる。その典型例が、it である。代名詞 it は話題として現れたものを指すので、it を多用すると理解に負担を強いることが多くなるようだ。原典では it が多用されているのに対し、書き直し版では少数である。

(12) 文長の効果を調べるため、単語ひとつを * ひとつで示し、センテンスごとの文長を調べた。次は原典の冒頭の 10 センテンスを例として示したものである。

```
 1 ******
 2 ******************************
 3 *********************
 4 ********
 5 ******************************
 6 ************************************
***
 7 *******
 8 ****
 9 ************************************
10 **************************
```

次は書き直し版の同じく冒頭の 10 センテンスである。

```
 1 ****
 2 *************
 3 ****************
 4 ****************
 5 **********
 6 ********************
 7 ******************
 8 **********************
 9 *******
10 ****
```

文長については、上の例にみられるように、また (9) に示したように、書き直し版では原典に比べ、平均文長もその標準偏差も小さくなっている。これは、原典では長い文、短い文が混在していて、メリハリの聞いた文章になっているということを意味している。書き直し版でもそのような工夫はみられるものの、そもそも、文長が長くないため、緩急のメリハリに弱みがある。
　(1)で、Flesch のリーダビリティー公式に関連して、*Reader's Digest* は *New Yorke*r よりも読みやすいと判定されるけれども、教育を受けた人の間では *Reader's Digest* は退屈で、*New Yorker* の方が 10 倍も読みやすいと考えたことを紹介した。この理由も上に述べた、文長とテキストのメリハリという点を考えれば納得がいくだろう。

(13) 次は原典と書き直し版のサンプルにおける 2 語連結の上位 20 位である。

of the	in the
in the	the Indians
to the	of the
the island	to the
and spices	on the
and the	to Spain
by the	and his
five hundred	at the
gold fields	Columbus and
he had	Columbus s
island in	for gold
it was	island in
on the	Spain with

the gold	that he
the Indians	the Arawaks
the slaves	the gold
to get	the Spaniards
all the	an island
an island	and more
an unknown	and queen

次は同じく原典と書き直し版のサンプルについて 3 語連結の上位 20 位を示したものである。

Asia it was	Columbus and his
gold and spices	to Spain with
island in the	and his men
to get a	at the fort
women and children	back to Spain
a battle with	back to Spain with
a battle with the	Columbus and his men
a certain quantity	in the Caribbean
a certain quantity of	island in the
a closer look	king and queen
a closer look at	that he could
a copper token	the king and
a copper token had	the king and queen
a developed agriculture	they had no
a developed agriculture of	to get a
a ferocious pace	A century later
a ferocious pace and	A century later no
a fort the	a closer look
a fort the first	a closer look at
a gold mask	a fort in

2 語連結、3 語連結とも統計的に明確な違いを見いだすことはむずかしかった。しかし、原典は書き直し版に比べ、総じて頻度が低い。おそらく、原典は表現のバラエティーが豊かであるのに対し、書き直し版では繰り返し現れる表現の割合が多い傾向を示すものといえよう。

(14) 英文テキストをやさしく、読みやすくする工夫をまとめてみると、次のようになる。

情報処理の単位を小さくする
　長い文を短くする。必要な情報に絞り込む。ひとつの文にいくつもの情報を盛り込まず、ひとつのトピックはひとつの文で表す。いわゆる「一義一文」である。
スキーマを豊富化させる
　背景知識のある大人であれば、必要な情報を補って理解することができる。子供の場合にはそのスキーマを補ってやる必要がある。この観点からやさしく書き直す工夫は「文体よりも冗長さを優先」「代名詞を使わず名詞で表現」「背景的情報を加える」とまとめることができる。
語彙を工夫する
　難語を避け、平易な語彙を使う。これは日本語では漢語よりも和語を使うことに似たものである。英語ではラテン系の語彙でなくゲルマン系の語彙を使うことで実現できる。
センテンスの構造をシンプルに
　とりわけ、関係代名詞の埋め込みを避けることが効果的である。

５．主な発表論文等
（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計１件）
朝尾幸次郎、立命館大学における新しい言語教育の試み、立命館高等教育研究、査読有、No. 10、2010、pp. 27-42

６．研究組織
(1)研究代表者
朝尾　幸次郎（ASAO KOJIRO）
立命館大学・文学部・教授
研究者番号：40102462