

# フランス語学習者スピーキングコーパスの構築とその分析—フィルターと中性代名詞を中心に—

杉山 香織

(東京外国語大学大学院博士後期課程 Acapela Group 研究員)

## 要 旨

コーパスに基づく研究は言語教育の分野に浸透しつつあるが、やはり研究の対象は英語教育の分野に限られている。本稿ではまず、学習者コーパスについて概観する。次に、フランス語学習者スピーキングコーパスの構築について説明する。最後に、初級学習者コーパスと母語話者コーパスとを比較し、使用語彙の頻度の差や他の語との共起関係を分析する。コーパス構築にあたって、ヨーロッパ共通参照枠の A1-A2 レベルの教材に基づき、学習者にタスクを与えた。このことによって、学習者の明確なレベル設定が可能となると共に、学習者から無理のない発話を引き出すことが可能になる。分析結果をフランス語教育現場へ還元することを最終目的とする。

## 1. はじめに

近年、言語学の分野においてコーパスに基づく研究の役割が大きくなってきている。コーパスに基づく研究は第一に、母語話者の言語使用の分析を可能にしてきた。このような研究は徐々に増え続けており、研究成果はさまざまな方面で評価されている。コーパスの有用性を受け、言語教育の分野でも同様に、学習者の言語使用を集めた学習者コーパスを分析対象として注目するようになった。しかしながら、学習者コーパス研究の大半はライティングに関するものが多く、さらには対象言語のほとんどは英語教育に集中している。フランス語学習者によるスピーキングコーパスはほとんど存在しない。日本人フランス語学習者によるスピーキングコーパスの存在に関しては、皆無に等しい。

フランス語は、日本において学習者の多い言語の一つである。また学習者にとって、フランス語の読み書きもさることながら、フランス語母語話者との会話もまた重要な学習動機の一つである。このような状況を考えれば、日本人フランス語学習者に特化したスピーキングコーパスが存在しないのはとても残念なことである。日本人フランス語学習者によるスピーキングコーパスを構築し、学習者の言語使用を分析することは、フランス語教育

の改善をはかる上で必要不可欠ではないか。

本稿ではまず、学習者コーパスに関して概観し、学習者コーパスの必要性和有用性について再考する。次に、筆者の構築したフランス語学習者スピーキングコーパスについて説明する。最後に、当コーパスを使用した分析と考察を行う。

## 2. 先行研究

### 2.1. 学習者コーパス

Sinclair(1996:4)によると、言語学におけるコーパスとは、言語サンプル収集時に明確に定められた言語的規定(または言語外規定)によって選定され編成された言語データである。今日存在するコーパスの大半は母語話者によるデータが中心である(Gavioli and Aston 2001: 240)。しかし、言語教育の分野に多大なインパクトを与えた学習者コーパスの存在も無視することができないであろう。

学習者コーパスとは、第二言語や外国語学習者による言語運用を、ある特定の目的に向け、明示的概念による規定によって収集されたデータである(Granger 2002:7, Granger 2003:465)。かつての言語研究の中心であったアプローチは、学習者による発話が自然か否かを母語話者の勘によって決定するものであった(Scott and Tribble 2006:3)。学習者コーパスは、従来言語教育で使用されたデータと以下の2点で大きく異なる(Granger 2003:465) : (1) 電子化されたデータ ; 幅広い言語処理ツールの使用が可能になり、言語データの分析をより緻密に行うことができるようになる (2)膨大なデータ量; 学習者の言語サンプルがより代表的となり、信頼性の高い結果が得られる。したがって、学習者中間言語の記述をより緻密に行うことができ、中間言語に影響を与える要因を明らかにできるようになる。さらには、研究成果を教育現場に活かすことで、学習者の要求に応じた教材や教室活動を可能にする(De Cook and Granger 2004 : 73)。

学習者コーパスは、分析の比較対照として母語話者コーパスを使用することが多い。母語話者との比較を通して学習者言語を研究する目的として、Tan(2005:127)は以下の3つを挙げている : (1)母語話者言語と比較して、学習者は目標言語に特有の語や構造がどのように不適切に使用されているのかを実証的データを用いて検証する(2)学習者の母語の特徴と、学習者のデータに特有の現象を関連付ける(3)分析結果をふまえて、教師が学習者に対してエラーに気づくような明示的な教育的フィードバックを与える。このように、母語話者の言語パフォーマンスと比較することによって、学習者の言語パフォーマンスの特徴が明らかになり、教育手法や教材開発への貢献をはじめとし、言語教育の分野に研究成果を還元することが期待できる。

学習者コーパスに基づく研究はこのように徐々に拡大され多様化されつつある。しかし、それは英語教育の分野が主となっており、フランス語に関する学習者コーパスは数が限られている。

## 2.2. 学習者スピーキングコーパスの有用性

学習者スピーキングコーパスに基づく研究は、ライティングコーパス研究と比べてまだまだ数が少ない。その理由として、データ収集、転写、また録音と転写等に必要なツールの開発などに、莫大な労力と時間を費やさなければならないからである(Luzón et al. 2007:3, Tono 2007:165)。しかしながら、スピーキングコーパスの重要性を指摘する声は多い(Compoy et Luzón 2007 : vii, Tono 2007 :164, Osborne 2007 : 181)。Tono (2007 : 176-177)は、スピーキングコーパスの教育現場への応用可能性として、(1)教材開発 (2)教師による評価方法の改善 (3)自律的学習の促進の3点を挙げている。特に、一つ目の教材開発に関して、学習者コーパスが担う役割は多大なものになるだろう。なぜなら Tono (2007 : 176-177) が指摘するように、ほとんどの市販教材は、著者の経験や勘、または従来の指導の型にとどまり、学習者の中間言語に関する学術的な分析結果を反映していないからである。Cheng (2007 : 199) や Biber and Reppen (2002 : 200) も同様に、教材で使用される表現は実際の言語使用を考慮しておらず、実証的研究に基づいていないものが多いと批判している。

学習者スピーキングコーパスの構築と、その分析は、教材開発をはじめとした外国語教育現場を革新する可能性を秘めている。

## 2.3. 学習者コーパスに基づく研究成果

コーパス言語学の特性のひとつとして、量的分析を挙げることができよう(Granger 2002: 4)。学習者言語に特有の語や表現を発見するには、コーパスに基づいた学習者言語の量的研究が不可欠である。母語、外国語に関わらず、いかなる言語サンプルを見ても、ある語やある連語がまたはある構造が、ほかのものよりも多い頻度で現れている(Bley-Vroman 2002 : 209)。つまり我々は、ある語や表現を繰り返す傾向にある(De Cook 2004 : 227)。しかし、母語話者によって使用される頻度の高いものは、学習者にとってもよく使用されるとは限らない。学習者言語は、質と量の観点で母語話者言語と異なっている：語、連語、構造の観点から、学習者によるそれらの過剰使用や過少使用の現象が有意な差をもって現れる(Granger 2003 :466)。また、学習者が、語や連語を過剰使用、過少使用、または誤用することによって、発話のぎこちなさや不適切な印象を聞き手に与える(Schmitt, Grandage and Adolphs 2004 : 139-140)。このように、頻度情報を使用することで、学習者言語の特徴を明らかにし、学習者の発話の外国語らしさの要因も明らかにできる。

コーパスの分析によって、コンコーダンサーを使用した共起語の研究も可能になった。Henry (1996: 296) は、共起使用に関する多くの研究成果はコーパス言語学によるものとしている。応用言語学の分野では長らく、語の共起関係について関心を向けてきた (Read and Nation 2004 : 23)。Granger (1998 :145) は、応用言語学において共起使用に関心を向ける理由として以下の点を挙げている:(1)Halliday や Sinclair によって提唱された語彙文法という概念は、語彙の連辞関係の研究によって進められてきた;(2) コーパス分析ツールの登場によって、語彙パターンに関する豊富で良質な情報を得られるようになった。

外国語教育において、特に共起語の指導は重要であり、目標言語への接触機会が少ない学習者にとってはなおさら有効である。なぜなら、共起語は効率よく記憶でき、また記憶

の維持もより可能になるため、学習者は日常生活におけるコミュニケーションに素早く対応できるようになるからである(Henry 1996 :296)。特に、共起語の学習は初級者に有益で、コミュニケーションの成功の第一歩を導く(Nattinger and DeCarrico 1992 : 183)。しかし、母語話者は簡単に共起単位を見つけことができ、また記憶することができるが、学習者はレベルによって共起単位を持つ量が異なっている(Spöttl and McCarthy 2004 : 191)。初級中級学習者言語の中に共起使用が見られないことも多く、共起使用を学習したとしてもその量が少ない(Osborne 2007 : 189-190)。そこで、母語話者コーパスと学習者コーパスを比較し、学習者によって過少使用されている共起語を明らかにすることで、学習者へそれらの共起関係を明示的に指導するのが有効であると考えられる。

### 3. フランス語学習者スピーキングコーパスの構築

#### 3.1. 日本におけるフランス語学習者スピーキングコーパスの存在意義

日本において英語学習者スピーキングコーパスは徐々に拡大され多様化されつつあるが、フランス語に関する学習者スピーキングコーパスは例を見ない。先行研究のとおり、学習者コーパスは言語教育の分野に多大な貢献をもたらしている。また、日本ではフランス語学習者の数は多く、コーパス研究がフランス語教育の分野で導入されないのは残念なことである。スピーキングコーパス構築には多くの困難点があるが、多くの学習者のニーズを満たすためにもフランス語スピーキングコーパスの構築が望まれる。

#### 3.2. CEFR について

フランス語学習の機会の大半は大学であり、そのほとんどの学習者が大学入学以降にフランス語を始めることから、日本で必要とされているフランス語学習者スピーキングコーパスは、初級者を対象としたものであるといえる。しかし、初級者といってもそのレベルはとても曖昧である。

そこで、より明確に学習者のレベルを規定するため、『外国語の学習、教授、評価のためのヨーロッパ共通参照枠 (以下 CEFR)』によるレベルを参考にした。CEFR は言語教育のシラバスやカリキュラムのガイドラインや試験、教科書の向上のために一般的な基準を与える目的で、欧州評議会が打ち出した枠組みである(Council of Europe 2008 : 1)。ヨーロッパにおける学習者にとって意味をもつ学習行程の区間を基にしたところ、6つのレベルを適当とした。これは、古典的な初級、中級、上級をさらに2レベルに分けたものと考えられる。CEFR では、この大きなレベルを A (基礎段階の言語使用者)、B (自立した言語使用者)、C (熟達した言語使用者)の3つにわけ、それをさらに枝分かれさせる『枝分かれ方式』を採用している(Council of Europe 2008 : 23)。熟達度の低い方から、A1, A2, B1, B2, C1, C2 となっている。

#### 3.3. フランス語学習者スピーキングコーパスの概要

東京外国語大学フランス語専攻の2年生である39人(男子学生10人、女子学生29人)

の協力を得て、2007年4月から7月までのフランス語の授業の間にフランス語のスピーキングタスクを行った。タスクはCEFRの基準に基づく以下の教材から、A1レベルのスピーキングに関するタスクを選んだ。

- (1) Beacco, M et Beacco, J-C. (2007) *Alors ? Niveau A1*. Didier ;
- (2) Breton, G., Cerdan, M. et Dayez, Y. (2005) *Réussir le DELF Niveau A1 du cadre européen commun de référence*. Didier ;
- (3) Crepieux, G. et al. (2006) *Spirale*. Hachette-Pearson Education Japan ;
- (4) Grandet, Eliane (2001) *DEL F A1/A2/A3/A4 : 450 activités*. CLE International ;
- (5) Parizet, M., Grandet, E., Corsain., M. (2005) *Activités pour le cadre européen commun de référence : Niveau A1*. CLE International.

表 1. タスク例

普段どのような映画を見ますか？
田舎と都会ではどちらの方が好きですか？
普段どのようなテレビ番組を見ますか？
朝食は何を食べますか？
好きな季節は何ですか？またその理由を述べてください。

授業終了後に行ったアンケート結果より、大半の学生は大学入学以降にフランス語の学習を始めており、長期フランス語圏滞在者はいないことがわかった。また、タスクの難易度を問う項目には、『どちらともいえない』と『やや難しい』に回答が集まっており、学生のフランス語能力はほぼ一定で、タスクの難易度も妥当だといえる。

録音には、東京外国語大学で開発した e-learning 教材システムの一部を使用した。学習者はタスクの指示を読んだり聞いたりして自身の発話を録音し、その記録はサーバーに保存される。転写はまず、学習者が録音を聞いて行い、後に筆者が録音と転写の一致を確認した。ただし、修正にあたっては以下の点に注意し、一貫性を持たせた：(1) 綴りや活用の間違ひは、発音が正しい場合、訂正する(例：Je m'apele\* -> Je m'appelle) (2) 発音上に違いがなく、転写時に性数一致を間違えている場合も同様に訂正する(例：bon\* restaurants -> bons restaurants) (3) 発話が不明瞭な場合や意味が通じない場合は、学習者による転写を優先する (4) 学習者が自身の発話の間違ひに気づき、転写時に修正した場合は、録音時の発話の方を優先する(例：[bo]と発音しているのにもかかわらず、belle と転写しており、後者が正しい場合でも、bon もしくは bons に書き換える)。

## 4. コーパス分析

### 4.1. 分析方法

新しく構築したフランス語学習者スピーキングコーパスを用いて、日本語学習者による過少使用を分析する。語や連語の過剰使用はコーパスの特性に左右されやすい。Shirato and Stapleton (2007:209)によると、過少使用の分析はコーパスの影響を受けにくいとしている。今回用いた参照コーパスとしての母語話者コーパスと学習者コーパスは、同じタスクに基づいておらず、サイズも異なるため、過少使用の分析のみを行った。分析には、*WordSmith Tools version 4.0*を使用した。*WordSmith*は、基となるコーパスと参照コーパスの語の頻度を比較することによって、過剰使用や過少使用されている「key word」を抽出できる。また、「key word」とともに、過剰使用や過少使用の度合いを示す「key-ness」の値が統計処理によって算出される。この「key-ness」の機能を用いて、過少使用を分析していく。

### 4.2. 参照コーパス<sup>1</sup>

参照コーパスには、ベルギーの中等教育最高学年である19歳から21歳までの学生の発話を集めたILRスピーキングコーパスを使用した。参加者は、勉強、旅行、趣味などの一般的な話題について、各フランス語圏におけるそれぞれのフランス語のステレオタイプについての2点をインタビュー形式で行った。そのうちの45インタビューをランダムに選択し、参照コーパスとした。

### 4.3. 両コーパスの生起総数

学習者コーパスと参照コーパスの総語数と異なり語数は以下のとおりである(表1参照)。2つのコーパスサイズには大きな隔たりがあるため、分析時にその点を留意する必要がある。

表 2. 生起総数

	総語数	異なり語数
学習者コーパス	28 602	1 920
参照コーパス	370 860	9 630

### 4.4. 日本人学習者による過少使用

日本人学習者によって過少使用されている語は以下のとおりである(表2参照)。今回の分析では、フィラー、中性代名詞の過少使用を中心に考察を進める。

<sup>1</sup> 詳細は Francard et al.(1993)を参照のこと

表 3. 学習者による過少使用の上位 50 語

N	Key word	NNS Freq.	%	NS Freq.	NS %	Keyness
1	EUH	48	0,17	7310	1,97	-781,48
2	ÇA	86	0,30	7616	2,05	-661,17
3	OUI	54	0,19	5187	1,40	-466,02
4	ON	96	0,33	6255	1,69	-454,32
5	QUOI	2	0,01	2905	0,78	-407,01
6	C	239	0,82	8713	2,35	-368,07
7	TOUT	14	0,05	2782	0,75	-316,99
8	ILS	9	0,03	2376	0,64	-287,20
9	ENFIN	4	0,01	1983	0,53	-262,31
10	QUI	74	0,26	4046	1,09	-257,34
11	MÊME	5	0,02	1894	0,51	-242,38
12	BIEN	57	0,20	3371	0,91	-227,54
13	OU	32	0,11	2435	0,66	-193,36
14	QUAND	29	0,10	2314	0,62	-188,64
15	LÀ	14	0,05	1657	0,45	-161,35
16	LES	166	0,57	4936	1,33	-151,64
17	COMME	22	0,08	1753	0,47	-142,66
18	SI	24	0,08	1738	0,47	-133,90
19	SAIS	14	0,05	1449	0,39	-133,70
20	PARLER	7	0,02	1215	0,33	-133,50
21	PARLE	7	0,02	1204	0,32	-131,97
22	ALORS	24	0,08	1707	0,46	-130,08
23	BON	33	0,11	1852	0,50	-119,79
24	PLUS	42	0,14	2038	0,55	-116,50
25	PEU	8	0,03	1103	0,30	-113,20
26	PEUT	9	0,03	1134	0,31	-112,88
27	PAS	331	1,14	7070	1,91	-98,67
28	DÉJÀ	4	0,01	839	0,23	-96,52
29	EN	182	0,63	4472	1,21	-92,29
30	DANS	60	0,21	2197	0,59	-92,14
31	ÊTRE	11	0,04	1036	0,28	-91,61
32	ONT	4	0,01	774	0,21	-87,37
33	Y	139	0,48	3599	0,97	-83,78
34	AUSSI	28	0,10	1395	0,38	-81,58
35	VRAIMENT	6	0,02	777	0,21	-78,04
36	GENS	14	0,05	1008	0,27	-77,29
37	MOINS	4	0,01	696	0,19	-76,46
38	NON	100	0,35	2781	0,75	-75,06
39	AS	5	0,02	704	0,19	-72,70
40	PAR	29	0,10	1335	0,36	-72,48
41	FRANÇAIS	51	0,18	1768	0,48	-68,77
42	UN	240	0,83	5027	1,35	-65,38
43	DES	223	0,77	4713	1,27	-63,15
44	DIT	7	0,02	677	0,18	-60,56
45	MAIS	179	0,62	3914	1,05	-58,45
46	TROUVE	6	0,02	595	0,16	-53,82
47	COMMENT	6	0,02	581	0,16	-51,99
48	QU	139	0,48	3119	0,84	-50,28
49	ICI	10	0,03	669	0,18	-48,96
50	RIEN	5	0,02	523	0,14	-48,40

#### 4.4.1. フィラー

過少使用されている上位 50 語の中に、「*quoi*」や「*enfin*」といった、ディスコースマーカ―のフィラーが母語話者によく用いられていることがわかる。

##### 《 *quoi* 》

学習者は母語話者と比べて「*quoi*」の使用頻度が低い(学習者 0.01%, 母語話者 0.78%)。この 2 グループ間には質的観点においても差異が見られる。「*quoi*」は疑問形としての使用と、フィラーとしての使用が考えられるが、学習者は疑問形としてのみ「*quoi*」を用い、母語話者はディスコースマーカ―として「*quoi*」を使用することが多い (表 4 参照)。

表 4. « *quoi* » のコンコーダンスライン

学習者
<i>quoi</i>
<id0703> Tu veux <i>quoi</i> Fabien ?
<id0726> Ce plat a <i>quoi</i> ?
母語話者
<i>quoi</i> + changement du locuteur
< ilrDN1> ça se voit ça ç/ se voit parfois <i>quoi</i>
< ilrDM2> c'est les gens qui se parlent euh // le moins bien le français <i>quoi</i>
< ilrBS1> pour euh pratique linguistique je mets bilingue mais en-fait c'est par l'école <i>quoi</i>
<i>quoi</i> + <i>mais</i>
< ilrDC5> je suis obligée <i>quoi mais</i> quelqu'un qui a une tête qui ne me revient pas je parle pas avec
< ilrDA1> oui il y en a un petit peu <i>quoi mais</i> oh je ne les connais pas vraiment
< ilrLD3> il faut pas pousser non plus <i>quoi mais</i> / il faut quand même un ce/ certain langage <i>quoi</i> / d'expression / politesse aussi ça c'est important

母語話者は「*quoi*」を用いることで、ひとつの談話単位を終えることが多い。参照コーパスでは「*quoi*」の後に話者が変わる使用例が 571 あった。また、「*mais*」と共に用いることで談話単位を一旦終えて、その後に逆説的な主題を始めるという用例が、166 例あった。つまり母語話者は、「*quoi*」によって談話を区切っていると考えられる。

CEFR のガイドラインにも、「*voilà quoi*」という表現は、自分の発言を言い終える機能を持つと紹介され、A1 レベルで習得するのが望ましいとされている(Beacco and Porquier



2007: 27)。母語話者コーパスでもその用例が 30 見られる(表 5 参照)。

表 5. 母語話者による《 *voilà quoi* 》のコンコーダンスライン

*voilà quoi*

< ilrVII > ça c'est curieux / *voilà quoi*

< ilrBM1 > je m'en fous parce que être Liégeois ou n'importe quoi enfin *voilà quoi*

< ilrLF1 > à la plage avec mon frère ou autre euh *voilà quoi*

しかしながら、学習者コーパス内にこれらの用例は一つも見られず、すべて疑問形として使用するのみであった。

#### 《 *enfin* 》

両コーパスで《 *enfin* 》の使用が見られるが、学習者は母語話者に比べこの表現を過少使用している(学習者 0.01%, 母語話者 0.53%)。詳細に見てみると、学習者は物事を列挙する際、「最後に」という意味で《 *enfin* 》を使用しているのに対し、母語話者はフィラーとして、主にためらいの表現と共に使用している。その例の一つである《 *euh enfin* 》は、参照コーパスの中に 101 例見られた(表 6 参照)。

表 6. 母語話者による《 *euh enfin* 》のコンコーダンスライン

*euh enfin*

< ilrPS1 > je trouve / *euh // enfin* ça dépend ce qu'ils mettent bien sûr

< ilrJO1 > je déplace de temps en temps à des endroits par *euh enfin* c'est juste aller au concert et revenir quoi

< ilrDC3 > ne pas bafouiller quand on parle / ou / pas faire de fautes *euh enfin* connaître beaucoup de vocabulaire // c'est bien parler le français

つまり母語話者は、言い換えをしたり、新しい談話単位を始めたりする際に《 *enfin* 》を使用するのを好んでいる。学習者コーパスではこの様な使用例は一つも見られなかった。また、《 *enfin* 》と似た機能を持つ《 *en fait* 》という表現は、母語話者コーパスでは 348 例見られたが、学習者コーパスでは 3 例しか見られなかった (表 7 参照)。

表 7. « en fait » のコンコーダンスライン

<i>en fait</i>	
学習者	
<id0702>	ah merci à votre votre gentillesse mais je n'aime je n'aime pas le foot <i>en fait</i> donc ah je suis désolé je ne peux pas aller avec avec toi
<id0715>	je suis désolée mais <i>en fait</i> je n'aime pas beaucoup le football
<id0733>	il est un peu méchant mais <i>en fait</i> il est très gentil
母語話者	
<ilrVS1>	ben c'est ça m'a <i>en fait</i> ça ça m'a aidé à à me décider
<ilrDS1>	en oral euh / <i>en fait</i> ce qu'il y a c'est que avant j'étais j'étais très ti/ très timide
<ilrbm2>	je ne sais pas <i>en fait en fait</i> si / si c'est la réalité ou pas on ne se voit pas

以上の分析を通して、学習者はフィラーの機能をあまり用いず、そのため談話単位を終える方法が不自然であったり、言い換えのストラテジーが欠如したりしていると結論付けることができるだろう。しかし、言い換えのストラテジーについては、CEFR の A1 レベルではなく、B1 レベルでの能力記述文に入っているため、A1 で求められるストラテジーではないといえる。

#### 4.4.2. 中性代名詞

学習者による過少使用のリストの中に、2つの中性代名詞 « en » と « y » が入っている。それぞれの中性代名詞について詳細に分析を進めていく。

##### « en »

学習者による « en » の使用は、全体の 0.63% なのに対し、母語話者の使用の割合は 1.21% である。「en」は、前置詞としての使用の場合もあるので、前置詞としての使用総数と中性代名詞としての使用総数を分けたところ、2つのグループ間で有意な差があった。以下の通りである(表 8 参照)。

表 8. 前置詞としての « en » と中性代名詞としての « en » の総数

	母語話者	学習者
en (前置詞)	3300	158
en (中性代名詞)	1172	24
合計	4472	182

$$\chi^2_{(1)}=15.527, p<.001$$

学習者コーパスでは、「qu'est-ce que tu en dis?」という表現が 8 例見られる。これは、友人と週末の約束をするというタスクに対する回答の一部である。使用数はあまり多くない

ものの、学習者による中性代名詞としての「en」の使用総数の3分の1であることを考えると、学習者はこの表現を固定表現として覚えていると言えるだろう。

一方、母語話者による「en」を含む表現では、「il y en a」(307例)が最多で見られた。学習者コーパスでは、この表現は1例見られたものの、正しい使用ではなかった(表9参照)。

表9. 「il y en a」のコンコーダンスライン

<i>il y en a</i> 学習者 <id0725> <i>il y en a</i> beaucoup de scènes horribles *  母語話者 <ilrGS1> je suis pas d'accord <i>il y en a</i> / beaucoup qui parlent wallon <ilrDC5> <i>il y en a</i> certains qu'il faut savoir se farcir quoi sans rien dire <ilrVI1> <i>il y en a</i> d'autres qui sont habillées tout à fait normalement
--

この差異は「y」の過少使用とも関係するので、「y」の過少使用を続いて分析する。

#### 「y」

中性代名詞の「y」も同様に、学習者は過少使用している(学習者0.48%, 母語話者0.96%)。「y」の使用方法として、大きく分けて2種類考えられる。「il y a」という定型表現に含む用法と、その他の中性代名詞としての用法である。「y」の使用法別に2グループ間で比較しても、有意な差が見られた(表10参照)。学習者は「il y a」という定型表現として「y」を使用することを、中性代名詞として使用するよりも好んでおり、その使用頻度差は有意であるといえる。

表10. 「y」の使用法別総数

	母語話者	学習者
Y	446	31
il y a (il n'y a pas, il y en a, etc.)	3153	108
合計	3599	139

( $\chi^2_{(1)}=11.806, p<.001$ )

今回使用した学習者コーパスと母語話者コーパスとも、この中性代名詞としての「y」は「aller」(行く)という動詞と共に使用されている例が多かった(表11参照)。

表 11. « y+aller »の時制別総数

	母語話者	学習者
<i>y + aller</i>		
不定詞	62	15
現在形	99	0
複合過去形	19	1
半過去形	9	0
合計	189/446	16/31

母語話者も学習者も一見同じ構文を使用しているように見えるが、時制別の使用数を比較してみると、母語話者の方がさまざまな時制を用いているのがわかる。学習者コーパスでは、不定詞 « aller » と共に使用する例を除き、ほかには複合過去との 1 例のみしか見られず、「j'y」と縮約されるべきところが « je y » と発音されており、使用に不自然さが残っている(表 12 参照)。

表 12. « y+aller »のコンコーダンスライン

<i>y + verbe « aller »</i>
学習者
<id0728> Je <i>y suis allée</i> avec une amie
母語話者
<ilrbm2> parce que elle elle fait des études à Louvain-la-Neuve / alors j'y <i>suis allée</i> quelquefois
<ilrBS1> j'y <i>vais</i> cinq fois la semaine quoi
<ilrCA2> il y avait une braderie à Ciney on on <i>y allait</i>

このように、中性代名詞 « en » と « y » は学習者によって過少使用されていることがわかった。また、学習者はそれらの中性代名詞を使用していても、中性代名詞を含む表現のバリエーションは少なく、質的に異なっているといえる。A1 レベルの学習者は、「qu'est-ce que tu en dis ? » や « il y a » といった定型表現以外は、会話において中性代名詞を使用するのは困難であると考えられる。

#### 4.5. 考察

今回は量的分析結果より、使用頻度の差が特に際立っていたフィラーと中性代名詞の過少使用に着目した。フィラーは特に、スピーキングに顕著な特徴であるが、今回対象とした学習者はスピーキングにおいてフィラーの使用には至っていない。また中性代名詞の使用に関しては、学習者は定型表現の一部として習得しているが、単独で使用するまでにはなっていない。

どちらの使用についても、A1 レベルの学生には少し高度なものかもしれない。しかし、

母語話者によるそれらの語や表現の使用頻度が高い以上、無視することはできないだろう。フィラーは日常生活でよく耳にするものの、教材ではあまり取り上げられない。学習者は、留学経験がなく、日本では母語話者との接触機会がほとんどない。そのため、フィラーのインプットの量が少ないといえる。一方、中性代名詞は文法事項に含まれており、習得済みであると考えられる。書き換え問題など、中性代名詞に関する文法知識を問う筆記問題には対処できる可能性は高い。しかし今回の分析により、口頭におけるアウトプットの過程には至っていないことが分かった。初級学習者における中性代名詞における知識とアウトプットの関連については追研究が必要である。

では、どのようにフィラーや中性代名詞のスピーキングにおける使用を学習者に指導することができるだろうか。やはり、母語話者による生きた会話のインプット量を増やすことが先決である。そして、母語話者によるこれらの語や表現の使用を、学習者が意識的に気付くように先導しなければならないだろう。そのためには、ただ単にインプットを与えるだけでなく、それらの項目を強調することが必要である。たとえば、今回使用した母語話者コーパスの中から、フィラーや中性代名詞が使用されているものを抜き出し、その会話を学習者に聞かせることによって、母語話者の発話スタイルに注意を向けることは可能である。日本のように、フランス語母語話者との接触が限られている学習環境では、母語話者のコーパスは生きた会話のサンプルとしても有効に利用できる。最終的には、自然な形で発話の中に現れるように、学習者にアウトプットの機会を多く与えるのが望ましい。アウトプットによって気付きが促進され、学習者はそれらの用法を習得する(Swain 2000: 99)。

## 5. まとめと今後の課題

本稿では、学習者コーパスについて概観し、フランス語学習者スピーキングコーパス構築の意義を唱えた。そして、実際に構築したフランス語学習者スピーキングコーパスの構築方法、内容について説明した。最後に、このコーパスを使用した分析例を紹介した。

今回使用した2つのコーパスは、同じタスクに基づいたものではないため、コーパスタイプに影響を受けにくい、過少使用に関する観点のみにしか言及できなかった。そのため、母語話者による同タスクに基づいたパラレルコーパスを構築することが望まれる。また、学習者コーパスについてもサイズがまだまだ小さく、もっと規模を大きくする必要がある。そして、コーパスに品詞情報タグなどを付加することも、学習者と母語話者の発話における構造的観点を分析する上で有効であると考えられる。

改善点は残るものの、本研究は日本人フランス語学習者の発話における言語使用の特徴をコーパスに基づいて行うという、新しい試みであった。その結果、学習者における過少使用の一部について明らかにすることができた。今後、コーパスの質と量を改善することで、より有効なデータを得ることができ、その結果をフランス語教育分野へ反映することによって、学習者のスピーキング能力向上への一助となるだろう。さらなる改善が期待される。

## 参考文献

- Council of Europe (2008) 『外国語の学習，教授，評価のためのヨーロッパ共通参照枠』 初版第2刷 吉島茂，大橋理枝（訳，編） 朝日出版社
- Beacco, R.C. and Porquier, R. (2007) *Niveau A1 pour le français – utilisateur/ apprenant élémentaire-*, Didier.
- Biber, D. et Reppen, R. (2002) “What does frequency have to do with grammar teaching ?”, *Studies in Second Language Acquisition* 24. Cambridge University Press: 199-208.
- Bley-Vroman, R. (2002) “Frequency in production, comprehension and acquisition”, *Studies in Second Language Acquisition* 24, Cambridge University Press: 209-213.
- Cheng, W. (2007) ““Sorry to interrupt, but...” Pedagogical Implications of a Spoken Corpus”, in Compoy, M. C. et Luzón, M. J. *Spoken Corpora in Applied Linguistics*, PETER LANG: 199-216.
- Compoy, M. C. et Luzón, M. J. (2007) *Spoken Corpora in Applied Linguistics*, PETER LANG.
- De Cock, S. (2004) “Preferred sequences of words in NS and NNS speech”, *Belgian Journal of English Language and Literatures (BELL) New Series* 2, 225-246.
- De Cock, S. et Granger, S. (2004) “Computer Learner Corpora and Monolingual Learners' Dictionaries: the Perfect Match”, in Teubert W. and Mahlberg M. (eds.) *The Corpus Approach to Lexicography. Special issue of Lexicographica* 20: 72-86.
- Francard, M., Lambert, J. et Masuy, F. (1993) *L'insécurité linguistique en communauté française de Belgique*, Communauté française Wallonie-Bruxelles.
- Gavioli, L. et Aston, G. (2001) “Enriching reality: language corpora in language pedagogy”, *ELT Journal Volume* 55/3, Oxford University Press: 238-246.
- Granger, S. (1998) “Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae”, in Cowie. A. P. (ed.) *Phraseology – Theory, Analysis, and Applications*, Oxford University Press: 145-160.
- Granger, S. (2002) “A Bird's-Eye View of Learner Corpus Research”, in Granger, S., Hung, J. et Petch-Tyson, S. (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, Benjamins: 3-33.
- Granger, S. (2003) “Error-tagged Learner Corpora and CALL: A Promising Synergy”, *CALICO Journal*, 20 (3): 465-480.
- Henry, A. (1996) “Natural Chunks of Language: Teaching Speech Through Speech”, *English for Specific Purposes, Vol. 15*: 295-309.
- Luzón, M. J. et al. (2007) “Spoken Corpora: New perspectives in Oral Language Use and Teaching”, in Compoy, M. C. et Luzón, M. J. *Spoken Corpora in Applied Linguistics*, PETER LANG: 3-32.
- Nattinger, J.R. et DeCarrico, J.S. (1992) *Lexical Phrases in Language Teaching*, Oxford University Press.
- Osborne, J. (2007) “Investigating L2 Fluency through Oral Learner Corpora”, in Campoy, M. C. and M. J. Luzón (eds.) *Spoken Corpora in Applied Linguistics*. Bern: Peter Lang: 181-197.
- Read, J. et Nation, P. (2004) “Measurement of formulaic sequences”, in Schmitt, N. (ed) *Formulaic*

*Sequences*, John Benjamins: 23-35

- Schmitt, N., Grandage, S. et Adolphs, S. (2004) "Are corpus-derived recurrent clusters psycholinguistically valid?", in Schmitt, N. (ed) *Formulaic Sequences*, John Benjamins: 127-151
- Scott, M. et Tribble, C.(2006) *Textual Patterns : Key Words and Corpus Analysis in Language Education*, John Benjamins.
- Sinclair, J. (1996) " Preliminary recommendations on Corpus Typology. Technical report", *EAGLES*.
- Spöttl, C. et McCarthy, M. (2004) "Comparing knowledge of formulaic sequences across L1, L2, L3, and L4", in Schmitt, N. (ed) *Formulaic Sequences*, John Benjamins: 191-225.
- Swain,M. (2000). "The output hypothesis and beyond: Mediating acquisition through collaborative dialogue", In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* : 97-114.
- Tan, M. (2005) "Authentic language or language errors ? Lessons from a learner corpus", *ELT Journal Volume 59/2*, Oxford University Press: 126-134.
- Tono, Y. (2007) "The Roles of Oral L2 Learner Corpora in Language Teaching : the Case of the NICT JLE Corpus", in Compy, M. C. et Luzón, M. J. *Spoken Corpora in Applied Linguistics*, PETER LANG: 163-180.