

1 コーパスから見える

フランス語

東京外国語大学 オープンアカデミー

秋廣尚恵

2 はじめに

- ・録音機機械の発明、その後の目覚ましい情報処理技術の向上により、大規模な言語データを蓄積、処理できるようになった。コーパス(資料体)の作成と応用は新たな段階を迎えている。コーパスは人文科学にとっても、社会科学にとっても、経験科学的な観点から重要なリソースを提供してくれる。
- ・様々な言語において、億単位の語数を持つ大規模コーパスが作成されている。フランス語コーパスの現状を紹介する。
- ・また、コーパスを使った学習法、調べものなど、身近な利用法を紹介し、様々な知識のリソースとしてコーパスを用いる方法を紹介する。

3 現代的な意味でのコーパスとは何か？

コーパス：資料体（ある研究目的のためにテキストを集めたもの）

→現代的なコーパスは、電子化されたサンプルとしてのデータである。

また、大規模なデータを機械処理するため様々なタグをつける場合がある。

- ・書き言葉では、書かれたテキストを電子化したもの。
- ・話し言葉では、録音されたデータを書き起こし、電子化したもの。

4 フランスにおけるコーパスのデータ共有の仕組み

- ・これまでフランスでは、個々の研究者、研究所単位でそれぞれの目的のためにコーパスを作成してきた。

- ・コーパスを一つにまとめ、共有しようという試みが始まった。

→ORTOLANG (Outil et Recherche pour un Traitement Optimisé de LANGue)

<https://www.ortolang.fr/market/corpora>

CNRTL (Centre National de Ressources Textuelles et Lexicales)

<http://www.cnrtl.fr/>

5 フランス語の書き言葉コーパス

- ・Frantexte <http://www.frantext.fr/>

16世紀から現代に至るまでの文学作品を中心とする書き言葉のコーパス。

語彙検索が可能である。

有料コーパス(一部無料公開の部分あり)

- ・Le Monde コーパス (European Language Research Association)

http://catalog.elra.info/product_info.php?products_id=438&language=fr

テキストファイルで新聞のテキストすべてが見られる。

6 フランス語の話し言葉コーパス

- Tcof (Traitement de Corpus Oraux en Français) ATILF

<http://www.cnrtl.fr/corpus/tcof/rechercher/index.php>

コーパスのタイプ、話者のカテゴリーなどに分けて検索できる。

- Eslo (Enquête sociolinguistique à Orléans) オルレアン大学

<http://eslo.huma-num.fr/>

Eslo 1 (1968~74) と Eslo 2 (2008~) によって構成される。

語数は既に800万語を越えるフランス最大の話し言葉コーパス

字列検索のみ可能。

-
-
-

7 フランス語の話し言葉コーパス

- TUFSS フランス語話し言葉コーパス 東京外国語大学

<http://cbile.tufts.ac.jp/tag/fr/index.php?menulang=en>

学生のインフォーマルな自由会話のみを集めたコーパス。規模は40万語ほど。

一部は品詞解析をしているので、以上のサイトから品詞検索が可能。

- パリ話し言葉コーパス <http://cipp2000.univ-paris3.fr/> パリ第3大学

パリに住むインフォーマントの会話を集めたコーパス

8 フランス語の話し言葉コーパス

- PFC (phonologie du français contemporaine)

<http://www.projet-pfc.net/>

書き言葉に関しては、量的に十分なコーパスが作成されているが、話し言葉コーパスに関しては、研究者ごとに異なる方針で作成されており、量的に十分なデータを持つものがない。→フランス特有の問題…

9 コーパスを使った研究成果

- 語彙研究として最も重要な成果！

T L F (Trésor de la langue française)

Frantext を基盤として作成されたフランス語大辞典 1994年に紙版が出来、

2000年以降は無料で電子版がインターネット公開。

<http://atilf.atilf.fr/>

10 コーパスを用いた研究の重要性

・現在はいかなる分野であれ、コーパスに基づいた研究は盛んである。

□音韻論的研究

□統語的研究

□談話的研究

□文体の研究

□文学作品の分析

□社会言語学的研究(借用、新語、話者による話し方の違い)

□方言研究

□フランス語圏の様々な多様性の研究

□

11 コーパスを使ってフランス語を調べてみよう！

・語法を調べる

例) après という言葉はどのように使われているだろうか。コーパスの中で観察し てみよう。

辞書には現れていないのに実はよく用いられている用法を知ることが出来る。

・ある特定のコーパスの中で用いられる言葉の特徴を知る。

例) ある作家の使う「色彩」を表す言葉。

ある作品に現れる「お菓子」を表す言葉。

受身形の使われ方をテキストごとに調べる。

12 コーパスを使ってフランス語を学んでみよう！

・文法書で規則を学ぶだけでは、フランス語が使えるようにならない！

→ 語彙と文法規則の結びつきが分からない。

→ 状況やコンテキストに応じた文法規則の使い方が分からない。

コーパスを活用する。

→ 語彙や形式を検索してみる。

→ 実際に話されている会話を観察する。

<https://www.lextutor.ca/conc/>

<https://apps.atilf.fr/fleuron/>

13 **コーパスを使ってフランス語の多様性をもっとよく知ろう！**

- 南仏で話されているフランス語にはどのような特徴があるのだろうか？
- カナダやアフリカのフランス語はどのような特徴があるのだろうか？
- 若者言葉、くだけた俗語ではどのような表現をするのだろうか？
- 政治家たちのスピーチで必ずよく使われる言葉はなんだろうか？
- 流行語、新語、借用語はフランス語の中にどのように定着、変化、あるいはすたれていくのか。

14 **フランス語の規範遵守の傾向**

フランス語は19世紀以降、学校教育の義務化、中央集権的なあり方を求める中で、とりわけ戦後、フランス語の多様性を切り捨ててきた経緯がある。

フランス語には現在、ダイグロシア的な状況が生じている。

- ・リングフランカとしての規範的なフランス語
- ・インフォーマルなコミュニティ内言語としてのフランス語
- ・

15 **インフォーマルなフランス語の研究から見えること**

- ・ラップのフランス語研究
- ・方言のフランス語研究

◆現代のフランス語において、大切なのは、様々なレパートリーを持った複層的な姿としてのフランス語を立ち上げること。

16 **データは客観的なものか？**

- ・コーパスを転写する段階で、あるいはコーパスとなるテキストを選ぶ段階で、実は私たちは主観的な選択を行っている。
- ・どのようなコーパスを用いるかによって、そこから得られるものは変わり得る。
- ・どのようなコーパスを用いているのか、常にデータの特性をよく吟味したうえで使用することが大事である。
- ・

17 **参考文献**

Landure, C. (2009). Data-driven learning, apprendre et enseigner à contre-courant,

Mélanges CRAPEL n°31 n. spécial : Des documents authentiques oraux aux corpus : questions d'apprentissage en didactique des langues, Nancy ; Atilf, p.163-178.

BILGER, M. (ed.) (2000). *Corpus, Méthodologie et applications linguistiques*, Paris ; Honoré Champion.

BLANCHE-BENVENISTE, C., ROUGET, C., SABIO, F. (2002) *Choix de textes de français parlé*, Paris ; Honoré Champion.

Sylvain Detey, Jacques Durand, Bernard Laks, Chantal Lyche (2010) *Les Variétés du Français Parle Dans l'Espace Francophone - Ressources pour l'Enseignement*