

平成 29 年度 東京外国語大学オープンアカデミー
東京外国語大学語学研究所 企画
『コーパスから見えることば・文化・社会』
2017 年 10 月 17 日 (火) 第 2 回
「コーパスから学ぶフランス語」
東京外国語大学准教授
秋廣 尚恵

皆様はじめまして。私は本学東京外国語大学でフランス語を教えております秋廣尚恵と申します。本日は皆様に、コーパスから見えるフランス語ということでお話をさせていただきます。どうぞよろしくお願ひ致します。

コーパスという言葉自体は前回成田先生が色々とお話になったと思います。少し繰り返すところもあると思いますが、コーパスとは何かというところから話していきたいと思います。

まず今回のお話で皆様に分かっていたかといいたいですか、伝えたいことをこの「はじめに」というところでまとめておきました。

録音機械の発明、その後の目覚ましい情報処理技術の向上により、大規模な言語のデータという物を蓄積して、処理することが出来るようになりました。

それで、コーパスというのが何かと言いますと、資料体と訳すことが出来ると思います。そしてその資料体の作成と応用というのは、実は今新しい段階を迎えています。昔は紙でただ書かれたものが蓄積されていたわけなんですけれども、もう今の時代というのは音であったり、マルチメディアであったり、またそれを扱うパソコンも情報処理の技術が発達していますので、新たな段階を迎えているんですね。コーパスは従いまして、人文科学——私たちは人文科学、外語大は人文科学の科目をやっていますけれども、そのほか社会科学にとっても非常に経験的な観点から大事なリソースを提供してくれます。つまり経験的と言うことは、何というんですか、頭でただ理屈を考えるだけではなくて、実際のデータを見て、それに基づいて、客観的に出来る限り客観的に分析していくということになります。そして前回成田先生からもお話があったと思いますが、ドイツ語も実は素晴らしいコーパスがたくさんありますが、今色々な言語においても本当に億単位の語数を持つ大規模コーパスが作成されています。

フランス語コーパスの現状を今回はまず紹介したいと思います。そしてみなさん、コーパスコーパスと言われて、コーパスにはこんなのあると言われても、自分の生活とあまり関係ないで終わってしまつてはつまらないので、コーパスを使って、例えばフランス語を勉強する方法にはどんな方法があるか、あるいはコーパスを使って、フランスの言葉だけではなくて、フランスの社会とか、フランスの文化などを調べていくことは出来ないか、そのような身近な利用法をここでは紹介したいと思います。そして、知識というところに書きましても、色々な意味での知識、ただの情報ではなくて、考えて自分のものとした知識のリソースとして、コーパスを使って、みなさんのフランス語、フランス文化の理解が少しでも深まればいいかなと私は考えています。

というわけで最初の方は恐らく成田先生もお話しになったことだと思いますけれども、簡単に復習の意味で、現代的な意味でのコーパスとは何か、ということからお話しします。

コーパスというのは先ほども申し上げましたように資料体——資料の集まったものですが、ただ集めればよいというわけではないんですね。あるそれなりの研究目的というのがあって、例えばバルザックの作品を集めて、バルザックの文体を調べたい。そのような目的があったら、バルザックの文学作品を徹底的に集めるわけです。そうしたように必ず何か一つ、研究目的をもって、そのためにテキストを集めたものをコーパス——資料体と言います。そして先程も言いましたように、現代的な意味でのコーパスというのは、まず電子化されている——つまりパソコンで見ることが出来る、検索できる、そしてまたサンプルとしてのデータである。これはとても大事ですね。

サンプルと申し上げたのはどういうことかということ、例えばバルザックの勉強をするのに、バルザックのたった一つの作品で、それがじゃあバルザック全体を語っているかということ、そういうわけではないと思うんですね。

ですからバルザック全体の作品を例えば代表するような、ある時代に跨った、あるいは複数の時代に跨った、そうしたものとある程度の量を集めないといけないわけです。

例えばフランス語の俗語——砕けた言葉を勉強したいと思ったら、コーパスというのはある程度一定の量を集めなければ、サンプルとして機能しないわけですね。サンプルというのは分かりますよね。化粧品でもいろんなもののサンプルがありますが、それが全体の代表として使われるということですね。ですからその全体の代表となるような質と量を備えたものでなければいけないというのが、現代的なコーパスの定義なんです。

また、ある程度大きな量のコーパスというのが扱えることになりますから、もういちいち手で読んで書き写してという作業では追い付かないわけで、そうすると何をするかといいますと、後でちょっと実際にお見せするんですけども、パソコンのソフトなどで一齐にこう検索をかける、というようなことが必要になるんですね。そのためには作業としていろいろなこのタグをつけるということがあります。

タグというのは何かということ、具体的に言うと例えば、一つの文章があったときに、これは主語ですとか、これは目的語ですとか、といちいちタグをつけていくのです。これは手作業だけではできないので、タグをつけるフリーソフトがあって大体それでタグを一回自動的にかけた後、後はもう手作業で間違いがないかというのを人間が直していくという、気の遠くなるような作業をやっていくんですけども、そのような形でタグを付けていきます。これが名詞だとか、これが何々だとか、情報を付けていくんですね。

それで実は、いまフランス語のコーパスを作成するのに、このタグをどうやってつけるかということで、研究者はものすごい激論を交わしています。これは非常に面白いんですが、私も何度かコーパスを作る現場の人たちと話をしましたが、物の見方によって一つの文をどう解析するかというのは変わってくるんですね。そうなると言語学者同士が——いい大人がですね、喧々諤々と喧嘩を始めるという、そういう感じなんですね。ですので、一ページの物をタグ付けしていくのに、二・三時間かけていくことも平気なんです。ものすごく気の遠くなる話だと思いませんか？

そしてこのコーパス——ちなみにこれは余談ですが、おいくらかかるかということ、フランスでは

一語作るのに一ユーロかかります。ですから一億語だと一億ユーロかかるんですね。すごいですよね。今、そういう気の遠くなる作業を、ちょっとずつ、言語学者が進めているという状況なんです。またそのための技術者というのも、今非常に足りていない状況です。ですので今後またちょっと大きな産業的にというか、非常に大切な分野になってくると思います。人的にも物的にも非常に高くコストのかかる分野です。

さて、コーパスには書かれたものと話されたものがあるんですが、書き言葉のコーパスというのは、書かれたテキストを電子化したもの。これはいいですよ。そして話し言葉は、それにプラスアルファ音が加わります。ですから、録音された音からまず始めないといけないんですね。話し言葉のコーパスはどうやって作るかというと、まず音を取りに行くんですね。話している人の録音をして、その次に音を書き起こします。書き起こしてさらに電子化をしていくという作業を取ります。

さて、これが大体いろいろな世界各国でのコーパスの様子だったのですが、ではフランスではどういうことが起こっているのかというのを、前回までの復習として、前回のドイツの話などとも比較しながら皆さんで聞いていただきたいと思います。

フランスというのはちょっと特殊な事情にあります。というのは、みなさん、フランスっていうと何となく、個人主義なんじゃないかとか、結構自分勝手な人が多いのではないかと思っていちゃいますよね。実際そうなんですね。言語学者もです。どうしてなんだろうと思うんですが、国民性といいますか、てんでばらばらなんですね。足並みが揃わない。これほど足並みがそろわず、仕事をしている人たちは居ないんじゃないかと傍から見ていて、何というんですか、ひやひやするといいますか、またなんでそうなるんだろうと心配なくらい個々ばらばらてんでばらばらにやっているんですね。ですのでそれはちょっと大変だと言うことで、さすがにフランス人もそれはちょっとまずいんじゃないかという人たちがだんだん出てきて、今まではもう個々の研究者、研究者単位で、とにかく好き勝手にやっていたんですが、これを一つにまとめようじゃないかという試みが生まれています。

それが実はこの ORTOLANG というものです。これは実際にインターネットのサイトにアクセスできます。この ORTOLANG というのは何かというと、**Outils et Recherche pour un Traitement Optimisé de LANGue** なんですが、言語の非常に最適化された処理のための道具と、それから研究であるという風に言っています。これは非常に大きな試みで、何をしているかという、この ORTOLANG のページに行ってみるとやはりこの個人主義のフランス人にはやりやすいように仕事を作ってくれているんですけども、それぞれの各個人が持っているコーパスをアップロードできるんです。

ただしそのアップロードに関して、相談に乗ってくれる総括をしている ORTOLANG の研究者のトップの人たちがいて、そのアップをするにあたって、ある程度の基準を満たしてないといけない。そう決めてくれているんですね。こういう基準を満たしてやってください。転写はこうでこうで。そしてその基準を満たせば、誰でもアップロードできるようにしてある。

そして私でも出来るんですが、私も相談に行ったんですが「君のコーパスはもうちょっと見直ししてください」と言われて、まだ見直しのためにはフランス人の手が必要なので、まだできません。ですが、私のようなものでもアップロードできるんです。そして素晴らしいことに、個々のサイト

というのは登録制で、研究者であればアップロードされたデータはアクセス可能になります。ですから無料で、ここでアップロードされたものというのはアクセスできるんですね。これは非常に素晴らしいシステムなんです。

そしてその一部が、実はこの CNRTL (Centre National de Ressources Textuelles Lexicales) という、これはフランスの国立研究所の一つの機関なんですけど、ここでもアップロードされていて、こちらの方は登録しなくても、みなさんが例えばご自宅から個々のサイトに行くと、全部ではないですが見ることができます。ortolangの方がかなり大きいですがこちらも大きいです。そしてこちらの方に重要なものいくつか、大切なものというのがもう無料で誰が見てもアクセスできるような形で置かれているんですね。ですので、もしフランス語のそういうコーパスに行ってみたいという人は、むしろ下の方のコーパスの方がアクセスは楽だと思います。

話し言葉と書き言葉に分けて話をしましたが、フランス語の書き言葉のコーパスというのは、これは非常に良くまとまっています。さすがにというのは、今までやはり文学の歴史が、滔々とした歴史がフランス文学はありますので、もうすでにテキストはある。やっぱりそれはあるんですね。印刷しているわけですから、活字にはなっているわけですから、そういうものは蓄積があるので、それをコーパスとして公開するという事は比較的容易にできるんです。

ここに挙げた FRANTEXT というのも、実はサイトがあって、ここのアドレスに行くとも見ることが出来ますが、残念ながらこれは有料コーパスになります。無料で公開されているのは僅か一部です。一年間の購読、要するに定期購読のような形で払っていくんですが、お金が大体一万円位だと思います。ちょっとお高いですよ。

東京外国語大学では、この大学内のネットワークでアクセスすることが出来れば、FRANTEXT は自由に検索できます。それは学生、それから先生方のために、大学が一年ごとのお金を払っているんですね。もしみなさん機会がありましたら、例えば外語大の図書館のパソコンなどから、FRANTEXT を見ることが可能です。

この FRANTEXT というのは、16 世紀から現代にいたるまでの文学作品を中心とするものなので、大体 8 割が文学作品、残り 2 割がその他の書き言葉です。例えば、先程言ったバルザックの作品だけ調べたければ、バルザックだけ取り出してみることが出来ます。そうではなくて、時代だけ、例えば 2010 年に出版されたコーパスだけ見たいとか、そういうことも可能です。あるいは 16 世紀のコーパスと 19 世紀のコーパスを比べてみたい。そういうのも可能です。データの幅も非常に広いですし、検索も割と可能に、またきめ細やかに作ってありますので、ぜひこれは試してみてください。

学生の中には色に関心のある学生がいます。例えばある作家の作品に出てくる赤という色のイメージを詳しく調べたいと思って、赤 rouge という色だけ全部抜き出して、コンテキストを調べて、その作家がどんな意味で使ったのか調べることも可能なんですね。こんな風に、言語学だけではなくて、文学などでも使うことが出来る非常に重要なコーパスです。

文学作品だけで FRANTEXT はほぼ 80 パーセント決まってしまうので、どうしてジャーナリスト的な文に関してはコーパスが足りないんですね。ではどうしたらいいかというと、その次に挙げたコーパスです。これも有料なので高くてついてしまうんですが、Le Monde というのがありま

す。Le Monde は知っていますか？新聞ですね。Le Monde の新聞はインターネットでダイジェスト版はただで見られると思います。私も Le Monde は必ずチェックしているんですけども、その Le Monde の全文です。テキスト全文というものをこの European Language Research Association というところから売り出しているんです。これもまた残念ながら有料で、一年分買うと五・六万してしまうので、ちょっとお高いかなという感じですが、これも大学では購入はしています。私の研究室にあるので気になる人にはお見せすることが出来ます。これはカタログがあって、このリンクに行くと買えるというだけなんですけれども、この Le Monde のコーパスというのは非常に素晴らしくて、テキストファイルでその年の新聞が全て読めるということです。ものすごい膨大な量です。ただちょっとやはり、加工などに時間がかかったり著作権などもろもろあって、全部が全部公開されているわけではないと思います。一部やはり削られているところがあるはずです。また、すぐその年のものが見られるようにはなっていないんですね。

私が見た最新のものは 2012 年で、それは一応こちらで購入したのですが、まあちょっと多少時差がある。ただその新聞コーパスなどを使って、例えばですね、選挙の時にサルコジがどんな言葉を使ったかとか、それをセゴレーヌ＝ロワイヤルと比べたという過去の実験もあります。サルコジは *je veux dire*、割と *je, je, je* とアピールするのに対して、セゴレーヌ＝ロワイヤルの方はどっちかという、*la france* とか、*je vous propose* とか、ちょっと柔らかな表現を使っているというような比較を真面目な顔をしてやっている研究者もいるんですね。

で、フランスというのはこれまた面白くて——いま日本も選挙ですよ。皆様も色々聞いてどれにしようかという感じだと思いますが、フランスというのはすごく政治に対する意識が高いんです。めちゃくちゃ高い。日本の学生を見ていると政治の話をしている学生は見ないんですが、大学院のレベルでもですね。選挙が近いですが政治の話をしている学生なんてほとんどいないですね。たまにいても自分の意見は絶対に言わないで一般的なことで終わるんですが、フランス人というのは割とその辺は議論をします。政治のことを言うと人間関係が壊れるというのはないんですね。これは素晴らしいなと思います。どうしてなの？と昔歴史をやっている専門家の友人に聞いたら、「当たり前よ。基本的な人権は誰にでも保証されているんだから。」と言うんですね。だからすごいですね。それはもう確固たる——保証されているんだから、後はもう戦うべきものは戦うという、だから喧々諤々やるんですね。

そういったこともあって、実はその政治的な効果つまり Le Monde を利用して、政治的な、政治家が何を言ったかみたいなことは、言語学者が必死になって調べ上げているところがあるんです。調べられる人はそれで人物を凶られているようで、恐ろしい気もするんですが、実際そうしたことはかなり流行と言いますか、非常によくやるんですね。日本ではあまり考えられないですけどね。その辺はすごくよくあると思います。

私実際に知っている、今修士の学生でフランス人の学生がいますが、彼もまたサルコジの大統領演説——大統領の言ったことをコーパスにして彼の言葉を十分に分析しています。面白いですね。そんなこともありますので、実はこの Le Monde というのは非常に大切なんです。言葉から見えてくるものというのは意外に多いと思います。

そして次はフランス語の話し言葉についてちょっと見ていきますと、こちらの方は残念ながら――

—書き言葉というのは比較的統一は進んでいて、もちろん書かれたテキストというのはただ公開するかしないか、その著作権の問題を何とかするかとか、そういったことになるのでそれほど難しくはないですが——問題はこちら、話し言葉コーパスのほうです。

成田先生はドイツ語の方をお話しになっていてどうだったのか、その辺私もちょっとお聞きしたかったんですけども、フランス語の方はかなり遅れています。本当に遅れています。英語に比べても本当に遅れています。どうしてかという、先程も申し上げたように、個人個人の小さいコーパスはいっぱいあるんです。ただそれが眠ったままの状態、その人が自分の研究に使ったらあとはもうお蔵入りしてしまう。もったいないことにいいデータがどんどん前に出てこないということなんです。いくつかの大きな大学の研究グループの中では、そうしたデータというものを何とか一括して集めて、というような試みがあるんですが、問題はその大きな研究所同士が仲が良ければいいんですけども、その大きな研究所同士もあまり意見が合わなかったりすると、一緒に仕事が成り立たないんです。そうするといつまで経っても別々に仕事をしているので、残念ながら結局お互いに勢力の削ぎ合いをしているんじゃないかっていうような状況なのですね。さきほどお見せしました ORTOLANG の解決策として挙げられているんですが、そうしたもので少しでも変わっていただければいいと思います。

そうなるいくつか小さいコーパスを見ることになるんですが、一番目に挙げたのは Tcof (Traitement de Corpus Oraux en Français) という、これは atilf——日本で言う国語研究所みたいなものに相当します——が出しているコーパスです。これもこのサイトに行ってくださいと実際にコーパスをすぐ無料で検索できるんですが、コーパスのタイプ、例えば大人の人たちが話したものか、子供たちが話したものか、あるいはどんな条件で話したか、家の中で友達と親しく話したとか、家族と親しく話している様子なのか、あるいはビジネスの場で仕事の場で使っているのか、会議なのか、あるいは講演なのか、様々なタイプに分けて整理をしてあるんですね。ですからこれは非常にいいです。タイプごとに見ることが出来ます。

また話者のカテゴリーに分けて検索できるとか、年齢とか性別とか学歴とかを自分で指定して、それに合ったコーパスのものがちゃんと検索できるようになっています。そしてその検索していくコーパスを自分でかなりきっちり選べるので、その意味では非常に大きなデータですが、その都度自分の目的に合わせて使います。

例えばちょっと子供がどんな風にフランス語の *le la les* という冠詞を勉強しているのか、身につけているのかを調べたい。日本人は苦手だけど、フランス人の子供はどうなんだろう、と調べたいと思ったら、ここのデータの中の子供の部分だけ取り出して調べたり……。大きいデータではあるけれど、個々人の研究もちゃんとできるような感じでやっているんですね。これはまだちょっと語数が少ないです。だいたい 40~50 万語くらいだと思います。

というのは、このコーパスを作ってるグループというのが非常に慎重なグループで、このグループのは GARS——元々はエクスマルセイユ大学に拠点を持っていた、今もうないんですけども、話し言葉研究チームだったんですね。今はもうその研究チームは解体されて全国に広がっている状態ですが、彼らは文法的な分析に力を入れていたんですね。非常に慎重に分析をしているという感じなんです。ですからゆっくりゆっくり絞って推し進めていくんです。実際には 1970 年代からずっとやってはいたんですが、そういう意味で、中々ゆっくり、慎重だけど、遅々として進まないなあ

というようにやっているんです。ただデータの質は非常に高いと思います。分析とか転写——転写というのは書き起こしたものに間違いがないかということなんですけれど、そういう意味では非常にクオリティの高い、私は割と使うことが多いコーパスです。

その次の Eslo というのは、オルレアン大学の——オルレアンは分りますよね、パリのちょっと東側の、ロワールの方に行く場所ですけれども——オルレアン大学で作っているコーパスで、これはまた非常に有名なコーパスです。恐らく上の Tcof のエクスマルセイユのチームと同じくらいの時期にコーパスを取り始めたんですが、こちらのコーパスというのは時期が二つに分かれていて、68年から74年。これは実はイギリス人がとってたんですね。フランス語を学習させるためにデータとして必要だということで、イギリス人が取りに来た。で、イギリス人が転写した。色々問題があったので、フランス人がそのあとを引き継いでいるわけですけれども、そのあと少し間が空いて、2008年から Eslo2 というので、二つ目のコーパス——いまなお、拡充されている状態です。

このグループというのは、さっきの Tcof が文法に非常にこだわったのに対して、文法とかそういうのにあまりこだわらないんですね。ただ、社会的にこだわっているんですね。だからどちらかというと、年齢が違う人、性別が違う人、そういった話者のカテゴリーを変えていった時にどんな違いが出るんだろうか、新しい言葉というのはどんな風に生まれて、どんなふうに広がるんだろうか、借用語はどうだろうか——そういったことを社会との関係でむしろ見ていきたい、というチームなんですね。ですからちょっと転写はあまり質がよくなかったりとか、分析がどうかなというものもありますし、また検索も Tcof と違って、字列検索だけしかできなかったり、ということで、多少リミットはあるんです。ただデータの数が実は 800 万を超えているので——Tcof は 50 万くらいだと思うんですけども——フランス随一の大きさのコーパスです。しかもこれは無料で、確実に自宅からもアクセスすることが出来ますので、みなさんもこのサイトに行ってくださいと検索することが出来ます。音も聞けるので、どうやって発音しているんだろうというのを知りたいければ、スピーカーをクリックすると、その具体例の音を聞くことが出来ます。

さらに、これは自分の宣伝になってしまうんですが、私たちもささやかながらコーパスを作っています。まだ質がいまひとつなので ORTOLANG には載せていただけないんですけども、これは私と、私のちょっと上の先生である川口裕司先生と二人でモーターとなりまして、実際にフランスのパリ、ボルドー、エクスに行きまして、学生のインフォーマルな自由会話をとりあえず録って、それを転写して、一部はタグ付けをしているので、品詞で検索ができます。字列だけではなく、動詞だけとか、そういう形で検索もできるんです。これが規模 40 万なので、語数はなかなか頑張ったかなという感じですけど、まだまだ質を向上中です。でもみなさんがここに行くことが出来ますので、よければ是非見てみてください。

本当にインフォーマルなので、私たちはどうしてインフォーマルなものを取りに行ったかという、教科書で習うのはきちんとしたフランス語なんですね。巷で話されている本当の——何というか、自然なフランス語というのはなかなか聞けないし学生も分からないし、教員の方からしても分からないところが多いんです。でするのでそれをデータとして録って見てみようじゃないか——というのが一番の大きな動機なんです。そういうわけで、とりわけそのインフォーマルな、何というのかくだけたフランス語に焦点を絞って 40 万語集めてみました。

それからその次に、これは比較的新しくて2000年くらいからですけど、パリ第三大学——本学も実はパリ第三大学と提携しております、先方とも色々仕事をするんですけども、パリの話し言葉コーパスが出ています。これはパリに住むインフォーマントの会話を集めたコーパスです。

インフォーマントというのは要するに協力者です。実際録音に協力してくれる人です。で、パリに住むといっても色々出身が違いますので、必ずしもパリを体現するかというあたりは難しいところだと思いますが、これについても非常に細かくコーパスを指定して検索することが可能です。これも無料で公開されていますので、この二つ、ご覧になる機会がありましたら是非見ていただくといいと思います。

それから最後に、フランス語の話し言葉コーパスで忘れてはいけないのは、PFCと言われるもので、Phonologie du Français Contemporain、現代のフランス語の音韻論研究というものです。これは文法とかいうよりはむしろ音にこだわってるんですね。フランス語の音、しかもこの音というものは非常に多様な音を聞きたい、そういうことなんです。

例えばフランスでも、パリのフランス語と、地方のフランス語というのは音が違う。どんな風にも多様性が出るか、というのは面白いんですが、みなさん、カナダのフランス語と、フランスのフランス語というのは同じだと思いますか？例えばシンガポールの英語とイギリスの英語とは違いますね。本当に違いますね。もう全然音が違います。言葉も——語彙も全然違うんでしょうけれども、まず音で聞いて、これはちょっと同じフランス語に聞こえないでしょう、というくらい違うんです。そういったものを非常に細かく録音しています。

最初は、この projet de PFC というのは国内でやっていたんですね。でも最近では、これが国外にも出始めていて、国外のフランス語圏に行って、そういうフランス語を録るということをはじめています。ですからコーパスはどんどんどんどん大きくなっておりませんが、実は先々週ですか、プロジェクトの立役者みたいな人が日本に来ていました。私もお会いして、早稲田大学へ招待されて、早稲田大学でお話しをされたんですけども、今後の展望などを聞きますと、やはりフランス語というものが——これはみなさんも分かると思いますが、フランスのことだけではもうやっていけない時代になっていますよね。

フランス語圏というと、アフリカもあるし、アメリカにもあるし、色々ある。北アメリカにもあるし、オセアニアにもあるし、と。今その多様性というのを認めてあげないと大変なことになると、やっとなフランス人の言語学者が動き始めたところなんです。

というのは、日本語ではその人の言葉遣いというのは人を表しますから綺麗な言葉遣いをしようと気を付けます。フランス人ほど意識が高い国民というのは珍しい。例えば、ある程度皆様くらいの年齢になった大人の方というのはきちんと話すと思うのですが、私が常日頃学生と話をしていますとすごい言葉遣いなんですね。私もそんな言葉遣いをするんだなと思って聞いているんですけど、今日聞いた話でびっくりしました。「ほぼほぼ」というのを聞いたことありますか？みなさん「ほぼほぼ」と言いますか？そんなことを言っていると、小学生の私の子供が「個々の人が」と言うんです。なんだそれはと思って私もびっくりするんですけど、大体の人はというんですね。「ほぼほぼ」は大体という意味なんだそうです。でも「ほぼほぼ」というのは言うだけではなくレポートに書いてきてしまうという世界なので、ちょっとびっくりです。実はそういう風に、割と無頓着なんです。

結構学生同士も聞きますと、やっぱりどんどん分からない表現が出てきて、一年間フランスに行ってきたという学生にちょっと話を聞いてみたら、「一年間以内だけで分からない言葉が増えているんですよ、先生」と言うんです。ツイッターとか見ても分からない言葉が増えているんです。実は非常に自由であると同時に、規範的なものに対してリスペクトしようという、遵守しなくてはいけないみたいな義務感が、ちょっと薄い感じがするんですね。もちろん就職の面接に行ったら、きちんとしなくちゃいけないと思ってきちんとするでしょうけれども、それすら怪しい感じなんです。特にツイッターで書いてしまうので、話していることを書いているので、どこからどこまでが話し言葉で、どこから先が書かなきゃいけない言葉なのかというのはもう、教えないと大学生レベルでもがたがたになっている。

ところがフランスは違うんですね。フランスは小さいころから学校教育が義務化されて、何が起こったかという、フランス語の徹底教育が始まるわけです。特にこれは戦後大変なことになって、後でちょっとお話ししようと思ったんですが、今先に話します。

さっき言った、PFCの枠組みに取られた方言のコーパスの中に、こんな話が出てきます。

「戦後私は方言を話していた」——その人は南仏の人で、オック語も話せ、フランス語も話せる人です。

「ただで学校で方言を話したら、その日、何か変な青いバッジを付けさせられる」——つまり方言を話したら罰せられるんです。

話しちゃダメ。話したら、そこをもう一回フランス語で言い直しなさい、ということをやっていたんですね。それで今何が起こったかという、方言がいまフランス語では死に絶えていって、もうどんどん廃れてしまっていて、恐らく60代、退職したくらいの年代ではまだ残ってはいるけれども、それ以降の年代の人は習わないと話せない。私くらいの年代だともう、習わないと方言も話せない。

でも日本はまだそこまで行っていないと思うんですね。逆に方言から出た表現が——例えば「そなん？」とかいう、終助詞が「なん」になるのは、もともと関西かどこかの方言か——なぜか一世風靡してツイッターでみんなばらまかれて全国に広がってしまうような、逆にそのような現象が起きるくらいです。

方言がアイデンティティとして着せ替えのようにみんなにうけると嬉しいみたいな、そういうところが日本語にはあるんですが、フランス語は逆なんです。

もう徹底してフランス語で統一してやろうというそういう教育は、ある意味いい面に働いたところもあると思いますけれども、悪い面にも働いて、そういう多様性というものを消してしまった、という現実があります。実はそれがいいことなのか、そのような事態に今みんな気が向いていると言いますか、言語学者が寄与しているんですね。

テロリストの問題にしても何にしても、あまりにもフランスの教育というのは、上からが一つ押し付けて純化していくようなところがあって、そこから外れたものというのはもう全部削除されてしまう。市民権を持たなくなってしまう。実はそういう教育は間違っているんじゃないか——それを言語学者が言葉の教育についても間違っているんじゃないかと考え始めたのがこのプロジェクトの背景なんです。

このプロジェクトに関わっている人の一人に、早稲田大学のシルヴァン＝ドゥテさんという人が

いますが、彼が言うのは、こういうことなんです。フランス語というのは、非常に豊かなレパートリーを持っている。上は本当に素晴らしい、文学者のフランス語で、皆様が憧れるようなバルザックであり、モンテスキューであり、という様々な人がいるんですが、それとは別に、下のレベルというか、方言から——あるいは、皆様が家庭で話している言葉というものもあるわけですね。その下のレベルというのは、下だからという風に消してしまうことは豊かさを失うことなんじゃないか——そう考えるようになってきたんです。そこで、こういうプロジェクトを立てて、実際にはもうなかなか聞けないような方言を録ってみたり、あるいは方言というのは昔から、もちろん長い伝統があるのですが、何といたらいいか、フランス語の地域版というか、オック語ではなくて、オック語の影響を受けた地域版のフランス語というものを聞くことが出来るように、サイトを立ち上げてるんですね。ですからみなさんに是非見ていただきたいなと思います。

フランス語というのは、純化されて非常に高度なというか、もちろんそういうものもありますけれども、それだけではない。またそれだけで終わってしまったのは、フランスの豊かさというのは分らないということは、きっと見えてくるんじゃないかと思えます。

まあ実は紹介するような本というのは、私たちがいま翻訳しているので、翻訳が出たらそのうちそれも日本語で読めるようになりますので、読んでいただきたいなと思いますが、本当にフランスというのは、言葉の面でも豊かなレパートリーを持ってるんです。それは本当に皆様には、この講義を通して伝えたいと思います。

ちょっとここでコーパスのまとめをしておきますが、書き言葉に関して言いますと、量的に十分なコーパス——FRANTEXT だけで3億語くらいあります。

話し言葉コーパスについては、研究者ごとに異なる方針で作成されているので、実際に量的に十分なデータを持っているものは何一つないんです。

Eslo が 800 万語——これが一番だと思いますが、これはやはり先程もお話ししましたように、みんな勝手にやっちゃってしまっているところがあって、これはフランス特有の場合で、恐らく英語圏などではその辺はあまりないんじゃないかなと私は思います。

ドイツ語圏の方は、先週聞いた感じではどうでしたか？ 割とまとまっている感じでしたか——まとまってそうですね。語数のレベルがだいぶ違うんじゃないかなと思います。そのような事情があります。やはりコーパスもお国事情があるわけです。

ここから先は、少しコーパスを使った研究成果について見ていきたいと思えます

まず一つ目です。当然コーパスをやっている中で一番大切な成果というのは語彙の研究です。みなさん、言葉の研究において、この TLF (Trésor de la langue française) というのは本当に素晴らしい辞書です。フランス語をちょっとでも読める人がいたら、ぜひともこの TLF をよく見て読んでいただきたいなと思いますね。本当に非常に素晴らしい辞書で、これは実は先程紹介しました、FRANTEXT を基盤として作成されたフランス語の大辞典です。割と最近なんです。今までは様々、アカデミーフランセーズの辞書があったり、いろいろあったんですが、実は 1994 年に紙版が出来上がりまして——紙版は本学の図書館にもありますけれども、緑の、アルファベット順に並んで圧倒されるような何巻にも渡る辞書ですね。なんと素晴らしいことに、フランスの良いところもあるんです。先程

の基本的人権は守られているというのにも関連するんですけど、貧しい人にもこれを提供して、フランス語を勉強していただきたいという、そういう非常に何とか温かい配慮があるわけです。

教育に関しても、教育費はほとんどかからない。フランスはすごいです。もう教育費は全然かかりません。大学も私立の大学がないということもありますが公立の大学は無料です。とにかく教育に関しては熱を入れていまして、この何巻にも渡る辞書が無料でインターネットで公開されています。

これは素晴らしい——私も良く使っていて、学生にも是非これを使いなさい、素晴らしいんだからというので使わせているんですが、このサイトに行ってくださいと *Trésor de la langue française* という素晴らしい辞書に行くことが出来ます。素晴らしいです、本当に。

ただ一つだけ問題があります。この辞書というのは先程も申しましたように、FRANTEXT を基盤にしているので、書き言葉のデータでわかる範囲でしか書かれていないんですね。例えば話し言葉の例というのはあまり考慮されていないんです。ですから、この辞書を読んでもピンとこないときがあるんです。

例えば、*après* という前置詞があります。フランス語で「あと」という意味の前置詞です。この *après* というのは、話し言葉でも前置詞としても使えます。

après le repas 「食事のあと」

après que... 「～したあと」

節でも使う。そして副詞としても使う。

je viens après 「あとで来ます」

ところがもう一つ大事な用法が、話し言葉では沢山出ていて、それがいわゆる接続詞——単独で接続詞として働く例なんですね。例えば、

Je vais aller au cinema ce soir, tu viens avec moi?

「今日映画に行くんだけど一緒に行かない？」

Mais après, si tu veux pas.

「でもそのあとで君が嫌だっていうんだったらね」

というふうに、ちょっと付け足すような感じで、こういう発話を付け足すような *après* というのが結構あります。

で、実はこの TLF にはその *après* の記述が抜けてしまっているんです。というのはこのような *après* は、話し言葉でしか使われないんですね。みなさん、よく考えてください。日本語の「あと」、同じですね。私は自分の息子が言っているのでびっくりしたんですけども——びっくりしたっていうかその通りだと思ったんですが、うちの息子が母の日の手紙に書いたんですね。で、先生にばっちり直された箇所があるんです。それは何かというと、「いつも仕事してくれてありがとう。あとお弁当作ってくれてありがとう」。この時の「あと」というのは、話し言葉で使いますよね。こういう「あと」というのは、学生の話で「飲み会は武蔵境がいいな。あと駅に近いところね」——例えばこんな風に言うのが、コーパスで出てきますね。

実はこういう「あと」というのは、「あとで」という意味じゃなくて——つまり時間的にあとで起こることを表すのではなくて、ある一つのことを言ったらそれにちょっと付け足して、「あとね」という意味のつもりで使えるでしょう。実はこれと全く同じ用法が *après* にもあるんですが、この大辞

典には抜けているんです。えー残念、と思うんですが、まあ仕方がない。FRANTEXT がベースですから。

ですから、みなさん、ここで一つ理解していただきたいのは、データをどうとるかで分析が変わってくるということです。たとえ大きな辞書であっても、それが書き言葉をベースにして記述されたものであれば、そういうデータとして、そういうデータを使ったものとしての分析しか出てこないということなんですね。

ではその時に、この *après* はさっきも見ましたが、崩れたフランス語だから記述しなくていいんだ——などと思っていたら、そういうわけにいかないんです。というのは、これを実際に話し言葉コーパスで調べてみました——そうしたらものすごい数が出るんです。出ちゃうんですね。

日本語もそうでしょうか？今言った「あと」などと言うのは、みなさんそんなこと書かないですけども、言いますね。恐らくすごい数で一日のうち何回言ってるかなという。

今度、気になる人は数えてみてください。自分の奥さんが言ったとか、自分の子供が言ったという風に。絶対にあると思います。

そういうものをじゃあ切り捨てていいかという、そうじゃない。やっぱりそのためには、データというものを覚えてもう一回見る必要がある。やっぱりそれは大切なことなんですね。

コーパスを用いた研究というものの重要性というのは現代にはいろいろとありまして、今言ったのは TLF なんですけども……。

話し言葉に基づいた素晴らしい辞書です。でも実際にはそれだけでは留まりません。今はもうコーパスづいていいますか、特に言葉の研究に関してはコーパスなしには語れないという部分がありますね。言語学をちょっとかじった方はご存じかもしれませんが、20 世紀の智の偉人と呼ばれるアメリカの言語学者であるチョムスキーという人がいます。素晴らしい人です。ただそのチョムスキーという人はコーパスには大反対だった人なんですね。どうしてだか分かりますか？彼はこう言ったんです。「人間の頭の中には文法規則というものがあって、限られた規則で無限の発話をプロデュースできるという素晴らしい力」——創造性です。言葉の創造性。「だからコーパスなど言うのはつまらない。コーパスに出てこない例はいくらでも人間は作れる」というわけですね。

ところが、では彼がその後何をやったかという、結局生成文法というものです。結局生成文法の人やったことというのは、その内省的な、一人一人の話者の直観に訴えるわけなんですね。

この文は合っているか、合っていないか。で、さきほど言った「あと」の話もそう。*après* の話もそうですけれど、正しいか正しくないか、ということ进行测试していくわけなんですね。ところが大きな落とし穴があります。ある人が、例えば私が「ほぼほぼ」は嫌だと思ったら、「ほぼほぼ」を間違いだと言ったら、それは正しいことですか？よく考えてみてください。人の直観、言葉に対する直観というのは実はそういうことなんです。その人の美的感覚といえますか、非常に心理的な感覚、社会的な評価であるとかに寄ってしまうんですね。例えば私はこういう教壇に立って、「ほぼほぼの人が」なんて言ったらやっぱりおかしいなと思うわけです。

でも学生が言って聞くのはまあ何となく、聞いて分かっていたりするわけです。ですから、直感というものは一体何かと言った時に、ある文を見て正しいか正しくないか——合っているかあっていないかと言えるのは、恐らくは、その人の心理的な価値判断しかないということは分かりますよ

ね。ですから逆に言うと、そういう生成文法派の研究はあまり批判したくはないんですけど、結局は生成文法派の人がやっている研究を見ていて私が一番納得がいかないのは、ではこの研究で合っているとか合っていないとかというのは誰の直観を調べているんだろうと思うわけですね。その筆者の直観しか調べられないでしょう？でもそれでは科学的な根拠を成さないはずなんです。

だからここで登場するのがコーパスなんですね。それは非常に大きな進歩で、今は生成文法、フランスでも生成文法の人はいらっしゃるんですが、生成文法の人とはいえ、まずコーパスで調べてから、そしてはっきりしないところを多くの人に聞いて直観を調べる、というようなことをしていると思います。ですので、実際には現代では、いかなる分野であれ、コーパスに基づいた研究が盛んになっています。

音韻論的、音ですね。音のシステムの話。それから統語というのは文法です。そして談話というのはさきほど言ったディスコース。例えばあの政治家の談話の場合、この人はこういうことを言うときに、どういう手段で、どういう言葉遣いで言うかというようなことを研究したい。

あるいはこういう研究もあります。どういう風にフランス人というのは、作文する時に論理を組み立てて、接続詞はどうやって使って、どんなふう論拠を上げるかとか。こんなことを研究したい。

あるいは文体の研究。例えばバルザックの文体の研究、文学作品の分析。これも色々できます。

それからさきほど言った、社会言語学的なオルレアンがやっているような問題です。まさに話者によっての話し方の違いであるとか。

あるいはある程度の時間を通じてみた言葉の変化。とりわけ借用語。あるいは新しい言葉がどうやって入って変化して、また廃れていくかというようなことを研究していきます。

それから方言研究。方言研究に至っては、元々コーパスに頼ってやっていたところがあります。グルノーブル大学というところがありまして、そこの方言研究者の方が最近こちらの大学にお見えになって講演をされたときに、とても面白かったんですが、そうしたコーパスなども地図とともになんか電子化されているんです。そういう非常に膨大なコーパスの研究というのは、実は方言研究では昔から当たり前だそうです。やはり実地的に集めたものでないと、逆に言えば、直観が利かないわけですから方言研究のようなものは出来ないわけです。やはりこの方言研究というのは、元々コーパスが強い部分であります。

それからさっきもちろっと言いましたが、フランス語圏の様々な多様性の研究があります。

これは先程話したんですけども、コーパスを使ってフランス語を調べてみようというので——まあ一つはこれ皆さんが実際にどうしたらいいかというので、さっき après というのは拾ってしまったので、繰り返しになってしまうんですけども……。例えば先程紹介した、とりわけこの無料でアクセスできるものに関しては、実際に行っていただくことが出来ます。大体字列検索はどのコーパスでも出来ますので、字列検索をやっていただくと、一つの語法というものを詳しく調べることが可能です。勿論フランテキストで調べれば、書き言葉の用法が分かる。

先程言った話し言葉で調べたら——先程言いましたね、ちょっと書かないけれども、話す時にいっぱい出てくる、というような用法が出てきます。そうした用法というのを知ることが出来ます。

コンテキストを出してくれるので、そのコンテキストを読むことで用法が分かっていくわけなんです。

そして先程も見ましたように、ある作家を使う。例えばこれは文学作品とか、文化的なことを知る時に非常に使えると思いますが、色彩を表す言葉を調べてみたい。何というか、言葉というのは不思議ですよ。日本語で赤と言った時と、フランス語の **rouge** というのはやっぱり違うと思うんです。何か、同じ色を本当にさしているんだろうとか。その色にどんなイメージがあるのか——いいイメージなのか悪いイメージなのか。やっぱり知りたいですよ。やっぱりそういう時というのは、こういうコーパスを利用して、どんな風に使われているか、言葉が使われているコンテキストを見ることでわかることが多いんです。もちろん実際にそれだけじゃなくて、柄を見たりということももちろん大事なんです。ただコーパスを使って、言葉の方からアクセスすることも出来ますよ、ということです。それから、お菓子を表す言葉。どんなお菓子があって、とかね。面白いことはいっぱいありますね。

この間、先程も話しました、方言の先生と話していたんですが、まあ話をしてくれたんですが、「イタチ」に当たる言葉。これがとても面白かったんですね。「イタチ」は分かりますよね。いろんなイタチ——フランス語では **la belette** というんですね。すごく迷信的な信仰の中では、とても大切にされる動物なんです。どうしてかという、お稲荷信仰じゃないんですけど、怖いんですね。元々害獣じゃないんですが、害を成すイタチというのは、とりわけ鳥を食べちゃったりとかね。もう一晩のうちに鳥小屋が全部全滅させられるような——すごくちっちゃな動物なんですけど、害をなす。だから農民の人たちというのは怖いんですね。あんなキリスト教の強い国でなぜこんなに、と思うんですが、それが民間信仰の中では、「イタチ」——キツネ信仰じゃないんですけど、**la belette** 進行というのがあって、習慣で、お鍋になんか餌をあげて、お供えしたりってこともあるそうなんです。その分布が言語地図だとか痕跡からわかるそうなんです。面白いですよ。カトリックの国なのに、実は日本人がとても理解と手も出来そうな、そういう民間信仰というのが知らないところであるんですね。そういうのは実は方言の言葉の研究から見えてくることがあるんですね。面白いと思います。

あるいはちょっとこれは学生向けなんですけれど、言語学の学生の、ちょっと頭の堅そうな学生が、受け身の使い方をテキストごとに調べる。受身の使い方。受身と言っても色々フランス語ではあって、日本語だとレルとかラレルを付ければいいんですけど——つけばいいというだけではない、と怒られそうなんです——フランス語は受身の形が結構たくさんあるんですね。「être+過去分詞」というのもあるし、**se verbe** と言って「代名動詞」の形もあるし、「使役受身」みたいなものもあるし、**on** という主語をちょっと受身に代用してみたり——すごくバラエティに富んでいるんですね。ですからその受身というものが、テキストによってどうやって使われるかというのはもう、すごく形の上で区別出来て、とても面白いんです。日本人が意味的にしか区別しないところを、形で区別しているようなところがあると思います。そんなのを見ると、やはり日本語とフランス語の違いというのが分かって、それはものの見方の違いであったり、あるいは物事を描写したりするときどんな手段を取るかということ——発想の違いでもあるので、二つの文化を突き合わせるうえで、とても面白い材料を与えてくれるんですね。なので学生は割とこうこういう研究なんかもやっています。実際うちの学校でアクセスして、そんな感じで皆さんが使えますので、なんでもいいと思うん

ですね。イタチ調べるのもすごくいいと思いますし、何か自分で気になったことがあったら、とりあえずその単語を打ち込んでみて、何が出てくるのかなというのを見てみるっていうのは面白いと思います。本で読むことというのは、確かに論理的に筋が通って、そして納得させられることもあると思うんですね。

ただ、こうコーパスを使うことのメリットというのは、皆さんが自分の頭で考えることが出来ることだと思います。そしてその考えるというのは自由な活動なので、やっぱりそういうものというのは、日頃考える癖というのは、とても大事なんじゃないかなと私は思います。色んなことを私も考えるんですけど、例えば言語学と言っても色々なその言語学者が素晴らしいことを色々言うんですが、納得できないことというのも多々あるんですね。だから言語学者は今でも戦ってるわけ——色んな人が。それで何かもうちょっと、そういう術語とか難しいだけのロジックは置いておいて本質的なことを語りませんか都下、思うこともあるくらい、たまに何か宙に浮いた、机上の論議などをしているわけなんですけども、やっぱりそういうのというのはある意味批判すべきところもあるわけです。その時に皆さん何をもって批判したらいいかというのは、まあこういうコーパス——自分がデータを見つけて、自分の頭で考えて、納得できるか、できないかということだと思うんです。納得できなければ、チョムスキーが言ったことだって、私はふーってやっていいと思うんですね。別にチョムスキー批判しまくっても別に誰を——ソシュールももちろん批判してくれても構わないだろうと思いますし、皆さんやはり自分でデータをもってきて、何というんだろう——知識というのはそういうことだと思いますね。何かを考える、自分なりに考えていくという先に見えるものじゃないかなと思います。そして考えていくときに、やはりその今まで先達が言ってきたこと、というのを色々比較することがあると思うんですね。

チョムスキーはこう言った。

ソシュールはこう言った。

アントワーヌ＝メイエはこう言っている。

その時に、やはりどういう見方というのが、どういう状況で、どう正しいのかということ判断する手掛かりを与えてもらえると思うんですね。ですから皆さんも色々言語学に関心をもって、もしいろいろなことを知ってらっしゃるにしても、まずはやっぱり自分の手で、自分で検索をして、自分で用例を見て掴んでいくと言うことを是非ともやっていただきたいなと思います。

それからもう一つここに挙げたのは、コーパスを使って、じゃあフランス語を勉強してみようじゃないか。

皆さんフランス語は——勉強された方いらっしゃいますか？

それでフランス語を——勉強するというのを、何語に置き換えて下さっても構いません、英語でもね。英語でも素晴らしいコーパスがありますので。コーパスを使って勉強すること、文法書で規則を学んだだけでは言葉使えないんですね。それはその通りで、どうしてかという、一番最初に考えなくてはいけないのは、その規則というものと、語彙がどうやってマッチングしているか、ということなんです。あんまり奇抜な組み合わせというのは間違いになってしまう、ということがありますよね。やはりよく使われる組み合わせから勉強していく、というのはとても大切なんです。で、その時の結びつきを見せてくれるのはコーパスなんです。一つの言葉を検索して、周り

に何があるか、と調べるレベルです。また、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。これもよくあることなんですね。さきほどの受身の話じゃないけど、être+過去分詞というのは、オーラルではあまり使わないんですね。会話ではね。ちょっと硬い文章であるとか、項客観的に物事を見て、判断する場合に——書く場合に使われることが多いんですね。新聞であるとか、書き言葉には多いんですね。ところがオーラルではむしろ、人を主語にして、題名動詞の使役受身みたいな形で代用させてしまうことがあるんですね。なぜかという、人に関心があるわけですから、やっぱり人を主語にして持ってきた方が話しやすいんですね。そういうことが起こるわけです。だからそういうのも、ある意味コーパスを見ていると、あっ、と分かるわけなんですね。なるほど、じゃあ話し言葉だから、受け身もこういう形とこういう形があるから、こっちを使ったほうがいいんじゃないか、などと見えてくるところがあるんですね。そうした規則の使い方を知るためにも活用できます。

ここに挙げたリンク二つ——フランス語を勉強されている方には役に立つと思いますが、一つはカナダのサイトで、**lexutor** というのがあって——これ実は、フランス語だけじゃなくて、ここに行くと英語やドイツ語や他の言語のコンコードダンスを見ることができます。これをぜひ見ていただくと——字列検索くらいしかできないと思うんですけど、それでもかなり検索して、実際に形式の周りにあるもの、コンテキストを見ることができます。

それから下に書いたのは、**fleuron** とあるんですけど、先程話した **atilf**——フランス語の国研みたいなところですよ——が出しているサイトで、これがどっちかという、会話をビデオで見せてくれるんですが、そのビデオで見せる会話を場面ごとに割り振って書いてあるんですね。

場面ごとに、例えばパン屋で買い物するにはどうしたらいいか、ビザを請求する時にはどうしたらいいかという、会話集であったりとか、というのがモデル会話じゃなくても、ビデオでそのまま見られるんですね。あるいは大学の登録の場合とか。あと留学生同士が話している場面で、フランスでうまく生活していくにはどうしたらいいかみたいなことを話していたりします。

で、これは何かという、留学生向けなんですね。やはりその留学生というのは、ある意味、フランスでは先ほども言った、移民の非常に多きな最初の入り口で、割と入りやすいので、まず留学して、そのあとについてみたい感覚になることが多くて、結構その留学生をどううまく入れているかというのは、フランスにとっては非常に大切な部分ですね。

やっぱりそういうこともあって、言語学者がそこに乗り出していく。で、どういう風に言ったらいいのかというような、教育的な効果もあるものを作ってくれてるんですね。こんなものもありますので、ぜひ観察していただきたい。日本でも多分あるかもしれませんね。

コーパスを使ってフランス語の多様性をもっと知ろうということで、これはある意味、ちょっとあとで時間を見て、聞いてみたいと思うんですけども、さっき言ったレパトリーを増やそうということです。これは皆さん、もしフランス語をやってなかったとしても考えて欲しいと思います。皆さんが NHK で教えてもらっている文法を、機知と——ドイツ語フランス語会話みたいなのは大事なんですが、最初は、最初は大事なんですが、それだけじゃないというのを、どこかで知っていてほしいと思います。それがとても大切なことなんですね。なんか、フランス人の話してるフランス語って変なの、と思ってしまうんですよ。そうじゃないと、それはまずいですね。そうい

うフランス語もあるんだっていうので、逆にそこに発見をして、心を開けるようになってほしいなとは、私は学生によく言うんですね。実際南仏で話されてるフランス語というのは、全然特徴がある——違いがあるんです。もう聞いて分かります。カナダ、アフリカのフランス語も同じです。もう言葉も、音も違いますよね。また若者言葉、砕けた俗語、全然違いますね。

日本語でも全く同じことだと思います。例えば先程言った政治家のスピーチ、あるいは流行語、新語、借用語。こういったものというのはいっぱいあって、言葉というのは非常に、複雑でモザイクなんです。しかも皆さん考えてみてください。一人の人間として、皆さん言葉を使う。日本語でも。ジャック＝デュランという、先程のPFCのプロジェクトのリーダーみたいな人がいるんですが、一日のうちに、一人の人間というのが話し方のスタイルをいかにも変えてるんですね。そうですよね。朝起きて、おはようと言って家族に言っている言葉と、会社に行っておはようございますと言って仕事始めてから使う言葉と、全然違うわけです。お昼休みにちょっと砕けて。ですからコーパスを見ると分かるんですが、一人の人が違う形を使い分けたりするんですね。実はそういうことも考えると、非常に言葉というのは多様でしかない。もう多様性の中にしかないということが分かるんです。その辺をすごく、コーパスを見て知っていただきたいなと思います。

先程も話しましたが、フランス語というのは、19世紀以来学校教育が義務化されて、中央集権的なやり方を認める中で、とりわけ戦後、フランス語の多様性を切り捨ててきてしまったんですね。ですから、現在フランス語というのは、非常にダイグロシア的な状況というのがあります。どうということかという、規範的なフランス語が上を覆っているんです。リングフランカとしての。それは確かで、そしてこう話さなきゃいけないというのは、徹底的に教え込まれてるんですね。フランス人は、学校教育で。しかも平等に。先程のTLFがなぜ無料で公開されているかといえば、そういうフランス語を広めるためなわけですから。そういうのは非常に熱心に広めているんですね。

ところが、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というのは、光が当たってないという状況で、まあ完全に分化してる感じではあります。それが何というのか、良いことなのか、悪いことなのかという、私は大切なことはどっちでも持つこと、だと思っんですね。ではインフォーマルな言葉遣いをしている人が、「じゃあああいいよ。君たちもインフォーマルなフランス語でやって言ってごらん」と言ったら、就職の機会はありますか、みたいな。分かりますよね。大学行けるんですかね？ ということになりますよね。だからそうした子供たちに、平等に教育を、素晴らしい教育を与えるというのは大事なことなんです。その子供たちの幸せはあるので。

ですが、それとは平行して、じゃあそれ以外のものを切り捨ててしまって良いかという、そうじゃないんですね。そこを切り捨てていくと、そういうところからはみ出てしまった人たちがいる意味不満を募らせて、歪んでいく社会になるんじゃないですか？ 言語学者がはっきりあるコロックではっきり言ったんですね。今このテロリストとか、問題がある中で、今そこインフォーマルなフランス語に、コミュニティない言語としてフランス語よりも光を当てなくちゃいけない、と言った。そこまで考えているんですね。やっぱりそれはとても大切なことなんです。市民権を与えるということですね。フランスでは方言に対する市民権というのは限られています。

皆さん今、カタロニアの情報が今入ってますね。素晴らしくはないんですが——あれは逆なんですけども、あれすごいんですが、実はそのフランスの国境のスペイン側とフランス側では同じ方言

を使っていますね。このカタロニアの方はカタロニア語ですよ。そしてこっちの方はバスク語なんです。ちょっと地図があれなんですけども。

実はフランスとスペインというのは、国境一本で全然状況が変わってきていて、フランス側の例えばバスクの人の話を聞くと、学校教育は二言語で行われていない——つまりフランス語だけで行われていないんです。ところが反対側のバスクでは、もう当たり前。もう学校で二言語で教育するのが当たりまえだ——社会的にもちゃんと市民権を得ている。それどころか、その文化的な多様性を認める以上にもう独立したがつているわけですよ。それがあるのは恐らく、今バルセロナでカタロニア州がやっていることだと思うんですが、そのことは一つ認められれば、恐らくバスクも同じことをやるわけで、もういろんな方言とか、みんな独立したがつてしまうわけですね。その政治的な独立というのは、フランス側では恐らく目立っていない、とされています。

これも実はコーパスを私は聞いたときに、バスク地方の青年が話していたんですね。

バスク地方の青年が言うには、「フランスではまず、僕が大切に思うのは文化的な多様性を認めてくれることなんだ。独立——政治的な独立というんじゃない。文化的にまず認めてほしい」で、その願いはやはり、ある程度聞いてあげなくてはいけないんじゃないかな、というような気がします。勿論それは政治的なものであり、あちらの人たちがみんな勝ち取っていくものであり、まあ選挙で投票して政治を変えていく、ということなんだろうなと思います。ただ、そういう傾向というのは本当にあります。

もう一つは、まあ社会的に政治的に、非常になんて言うんですか、南仏の——南の国というのは貧しいので。逆に、スペインの側はバスクにしても、カタロニア州にしても、経済圏が強いですからね。もう独立してしまった方が、後を切り捨てられるのでいいという、実際にそういうなんかちょっとお金の、財政的な問題もあるんだと思います。まあフランスの方のバスクとか、カタロニアの方というのは、独立はしていけないので、逆にパリに依存しなければいけないというのはもちろんありますから、実際にそういう問題もあると思います。ただそういう状況というのはあるんだよ、ということは知っておいていただきたいなと思います。

それから最後に、インフォーマルなフランス語の研究というものについて見えることを見てくださいと、ラップというのは皆さん知っていますよね。ラップ音楽。これも今研究がすごく面白いことになっているんですね。それで、セルジ＝ポントワーズ大学の Pruvost という、語彙の研究をしている先生がいるんですが、彼が *mouv'* というラジオ番組に、毎回 *Doc Dico* という名で登場しているんですね。そしてラップに出てくる言葉を紹介しているんです。

ところが、このラップに出てくる言葉というのが、どうしようもない言葉ばかりじゃないのか、と馬鹿にして聞いていたら、違うんですね。実は中世の、昔の古い表現が民衆語の中に残っていて、その生き残りみたいな表現がいっぱいあるんですね。これびっくりするんですね。 *la dalle* という言葉があって、 *avoir la dalle* 「腹が減る」という意味なんですけども、それをラップでは言うんですね。若者ことばで腹が減るという、非常に砕けた——ちょっとはしたないくらいの表現だと言われているんですが、 *la dalle*、 *avoir la dalle* と言うのはそもそもどこから来たかということ、14世紀に、スカンディナヴィア語から入ってきた言葉で、 *dæla* とかいう言葉らしいんですね。そして、その言葉が当時は、なんて言ったらいいか、水を流す——流しです。流しの穴の開いた部分。水がサーッと下に

流れて行く、この部分を **dalle** というんだそうなんです。そうして入って来たんですね。おかしいですよ。

ところが時代が変わるにつれて、その **dalle** というのは、現在フランスでは下に市区タイルみたいな意味になってしまっているのですが、元々その中世に——忠誠というか 14 世紀くらいに入ってきたときには、水を流して排水をする、排水溝の部分だったんだそうです。実はその意味が、ラップで使われている **la dalle** に残っているとされていて、つまり、「喉」がというんですね。そこから排水溝が流れる、というので、**rincer la dalle** という「喉を流す」という意味で、まあ「水を飲む」とか「飲み物を飲む」という俗語になり、**avoir la dalle** という——ここで言うのは食道ですよ。位に直結しているので、「お腹が空く」という意味なんだそうです。お腹が空いて、食べなくなるんですよ。食道から手が出るというか、食道を持つという意味で、**avoir la dalle** 「お腹がすく」。でもその起源は、実は中世にさかのぼってしまうという。そういう表現はいっぱいあるんですね。ラップの中でお金のことを **maille** というんですけど、**maille** というのも実は、中世の小さなお金の単位だったそうです。どうしてそういうことが起こってしまうかという、やはりその民衆語の中にこういうものが蓄積しているんです。恐らくね。忘れ去られないで、それが規範からずれたところで何となく残っていて、他の世紀になって使われるような言葉とは、ちょっと違ったカラーを持つ言葉として、ずっと使われてきたんじゃないか、と。だからラップというのは、必ずしも若者だけが知っている言葉じゃなくて、実はそういう民衆語から取ってきたものも——アラブ語からももちろん取ってきた言葉もあるし、いろんな外国語がから入ってきたり、あるいは方言から入ってきたり、昔々の表現から取ってきたり、というので構成されているそうなんです。そういうのを毎週一回くらい、ラジオ番組で流していると。若者は自分で使っているくせして、知らないんですね。で、わーっ、と喜ぶわけです。面白いと思うわけです。流行っているそうなんです。面白いですね。だから若者にとっても、そういう風にして、言語学で、インフォーマルで知られていなかったところに手をを入れて話してあげるといのは、とても大事なことなんじゃないかなと思います。

最後に、これはちょっとデータというものについて、もう一回行っておきますが、じゃあデータは客観的なのか。コーパスを使えばいいのかというところで、鋭くもう一度、批判的に——自己批判なんです、しておかなくてははいけません。

コーパスというのは、転写をしていく段階で——話し言葉の場合、あるいはコーパスをテキストとして選ぶ段階で、実は私たち、主観的な選択を行っている。書き起こす時点でどう描き起こすかというので、主観的な判断をしているんですね。ですから、実はデータというのは、何一つとして客観的なものっていうのは存在しえない。ここはやっぱり知っておくべきだと思います。ですからどのようなコーパスを使うか、どれくらいの量のコーパスを使うかということによって、そこから得られるものは変わってきてしまいます。そこはやっぱり非常に気をつけなくてははいけません。

ですからコーパス至上主義というのは、私はちょっと注意をしなければいけないと思います。コーパスさえ使っていればいい、というものではないということです。どん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わないといけない、というのを最後に言っておきたいとおもいます。これはまあ当然なんですよ。

経験的であろうと思うがためにコーパスを使うんですが、実際、何でもコーパス使えばいいかと

思って、適当に統計を取って、その統計の数字を弄って、提示をしても、納得できない研究というのはいっぱいあって、その多くは、やはりコーパスをよく吟味していない。コーパスを選んだことに対する、何というか、そのスタンスをはっきりさせていない、ということなんですね。恐らくデータというのは、多かれ少なかれ、理科系のことでも、社会学的なことでも、同じような性格を持つのではないかと思います。

というわけで、私の話は以上になります。(完)