

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 11 月 7 日 (火) 第 5 回

「アラブ映画のシナリオから見えたこと」

東京外国語大学非常勤講師

長渡 陽一

---

長渡です。今回担当いたします長渡と申します、よろしくお願いいたします。アラビア語のコーパスの話なんですけども、皆さんアラビア語というのをあまり知らないんじゃないかと予想して、いわゆるコーパスの話よりも今回掘ってお話ししようと思っています。

まずこの資料をご覧くださいと、皆さん、どの地域でアラビア語が話されているかご存知でしょうか。地図に上げておきましたが、1番東の方ではイラクとかオマーンとかアラビア半島の辺り、もちろんアラビア半島アラビア後の故郷ですから、ここは当然アラビア語を話します。そこからイスラム教と一緒に広がって、イラク・シリアなどアラビア語が広がったわけですが、エジプトがアラビア語を話す人が1番多く住んでいるところです。アフリカの中でも1番人口の多い所ですので、ここはかなり集中しています。それからリビア・アルジェリア・モロッコ・モーリタニアまでアラビア語が話している地域は結構広いですね、広いですが人が住んでいない地域、砂漠が広いので人口にしてみるとそれほど多くはないです。面積で言うと、エジプトが日本と同じ位です。

アラビア語を話す人口が大体3億人ぐらいですが、それも数え方によっていろいろです。話者数の数え方ってどういう風に数えるのでしょうか。まず1番問題になるのは、ゼロ歳児は日本語話者か？普通は日本語話者に入ります。話さないけど入ります。アラブ地域には他にも色々な事情があって、アラブ人ではない人が結構住んでいます。1番割合が多いのはモロッコで、モロッコは半分以上がアラブ人ではないです。モロッコでアラブ人以外の民族の名前は、どなたか、ご存知ですか？半分以上がベルベル人という。街に近くて学校に通えるような人は小学校でアラビア語を勉強します。街に出ればほぼアラビア語です。ベルベル語が半数と言いましたが、お互い通じないのです。なにになにベルベル語というのに分かれていて、ベルベル語が通じないので話す共通語はベルベル語ではないんです。そういうわけで、ベルベル語は有力言語とは言えないのです。街に出たらアラビア語だったりするわけです。そういう人をアラビア語話者に入れるかどうか？

また今問題になっていますが、クルド人はどうか？戦争になってISが活躍していますがクルド人が奪還してたというニュースがありました。あのような人たちはクルド語を話していますがアラビア語も話します、がクルド語で暮らしています。クルド語を話す人たちが難民で日本に来ましたけれども、じゃあ何語で日本語を教えたらいいか。クルド語で教えたなら良いのですが、クルド語の日本語の辞書がないのです。アラビア語でできるかというシリアのアラビア語は話せるけれども、書き言葉のアラビア語はできない。そういう人アラビア語話者、話せるからアラビア語話者に入れられるかというと、いろいろな問題が絡んでいます。今はヨーロッパにシリアから大量に渡っているの、シリアの地域ではアラビア語話者が相当減ってしまいました。

そう考えるとイエメンにもアラビア語ではない民族の人たちが住んでいます。エジプトのナイル川上流の方にはヌビア人と言う人たちが住んでいて、リビアの方には先程言ったベルベル人が住んでいるし、そういう人たちを除いたアラビア語を話す人たちが何人いるかについては統計がほぼ取れない。登録された人口が何人いるかと言うと、これは日本と違って正確な言語人口は出てこないのです。それを概算して、クルド人が何人ぐらいいるかをひいて、大体2、3億人と言うことになっています。広い地域で話されているので、差がかなり大きいのです。シリア・アラビア語とエジプト・アラビア語とは近いのですが、ベタベタのシリア弁の人とベタベタのエジプト人が傍で聞いていたら、半分わかるかなと言う感じです。その半分がどのくらいかというのは少しわかりませんが、単語的にはシリア・アラビア語とエジプト・アラビア語とはかなり近いのですが、発音の癖とかなまりがあるのでぱっと聞くとなかなかわからない。1番左に書いてある、...アラビア語、これはエジプト人が聞いたらず9割かたわからないでしょう。単語の使い方が違うので厳しい。

NHKでニュースをやっているイラクからのハーフの方とモロッコからの方と話すと、エジプトで話すアラビア語と文語の知識がありますので文語を入れつつ、なんとなく最大公約数のアラビア語で話しています。だんだん話してくると相手の様子がわかってくるかもしれない。それくらい話し言葉の方言差は大きいのです。

次に文語体と口語体。アラビア語の書き言葉と話し言葉の違いは大きいのです。今日本は書き言葉が分かると会話ならできますよね。今外国人は書いたものを読んでそれを勉強しつつ話し言葉に会話に応用できますよね。だから書き言葉と話し言葉の違いは、日本語はかなり近い。言っちゃうと言ってしまふ、は勉強すればかなりわかる程度です。ところがアラビア語はかなり違う。ではどういうところで使ってるかという、下のところで示してありますが、アラビア語ではアーミーヤと言って、会話は全てアーミーヤです。テレビから流れてくるアラビア語は文語体も使われる比率が高いのですが、基本的には口語体、アーミーヤです。友達と話すときにはもちろん口語体ですね。市役所の窓口に行ってもやっぱり口語体です。学校が問題ですが、先生や子供が話すのは口語体です。小学校に入って学ぶのは国語としてアラビア語として文語体を学びます。です。しかし文語は話せませんから解説等は全て口語体でします。では文語体はどういうところで使うか。本、新聞、メニュー、看板などは全て文語体です。文字で書かれているものは全て文語体です。ニュースは文語体です、これは原稿があるので、原稿書くときは全て文語体で書きます。国際会議、頻繁に開きますけれども、モロッコ、イラク、サウジの人々が話す時は、実際はかなり口語なのですがベースは文語体です。ただ文語体ですらすら全て話せる人は稀で、かなり教養が高い人です。口語体で話したいなあと思いつつ、これを文語体に置き換えて話しているのです。白熱してくるとだんだん口語体になってくるのです。文語体では話せないのです。

だからアラビア語は勉強しようとしたら、どうすればいいか、これは問題になります。まず勉強しようとなると文語体をやります。だからこの大学のアラビア語専攻の学生たちはまず1年間は文語を勉強します。ただし文語でするので会話はできないのです。会話の練習をするのですが、まずは文語体で会話の練習をしますが、現地に行っても使えないですね。文語で話していないのですから。だからと言って口語を先に勉強しようと言ってもなかなか難しいのは、文字にかかれませんか。外国語を勉強しようと言う時はまず本で勉強して読み物を読んで単語を覚えて、それを会話に使ってみようということになりますよね。だけれどもそれがうまくいかないのがアラビア語なのです。

だからアラビア語は難しいのです。後でお聞きいただきますが、アラビア語自体は全然難しくないです。ごく普通の言語ですからね。文字がふにゃふにゃとなっていますが、学生は大体1週間位で学びます。週一回の授業ですと、1回の授業でも読めるようになります。

文字は、見た目はちょっと抵抗がありますが、そんな難しくはないです。言語もそんなに難しくはないです。では何が難しいかと言うと、読んだり書いたりしたものを会話に応用できないと言うことです。では会話ができるから読み書きできるかというところもできない。それが別個なので2つやらなければいけない。それが非常に難しいところです。

ではその難しさを体験していただくために次の例を見ていただきます。日本語でちょっとやってみます。これは明治初期の明六雑誌と言うものですが、西周と言う人が書いた文です。これは日本語の文語体ですね。ちょっと読んでみますね1段落目からですね。《皮肉節と言う題名の文章です。

《かくの如き...》

これは文語体ですね。この雰囲気です。これで会話すると言うのは、相当な訓練が必要そうですね。書くのはいろいろなところから引っ張ってきて時間をかければ書けるかなあとと思いますが、読むのはなんとなくじっくり考えれば読んでわかるかなあと。でもこれで会話はできないですね。作文も結構大変ですね。私にとっては、生まれてから周りのお父さんお母さんなどの人の口語体の会話を聞いて学ぶわけですね。それで小学校に入ってこれを勉強するわけですね。読み書きがこれなんです。本を読むときはこれで読むのです。だから日本の大学の学生がこれで文語体を勉強した年とすると、それで夏休みになってエジプトに行ってみようとして、そこで会話するわけですね。

《かくの如き...》と。これで笑われて凹んで帰ってくるのです。なんか笑われちゃいましたと言う。それは文語体でしゃべるからです。外国人が文語体でしゃべったらちょっとおかしいですね。単語がいくつか違うのですが、同じものも結構あります。文法的には「掲げば」は「掲げれば」と現代では言うところですね。《いかんたる...》このようなのが現代のアラブ世界なのです。日本では明治時代の頃はこういう状態でしたけれども、今は書き言葉的に話し、書き言葉もかなり口語的になっています。口語的になって、ほぼ一致しています。それがアラブ世界では一致していないのです。それで日本のアラビア語学習者と言うのは結構困るわけですね。日本に限らずアメリカでもヨーロッパでもアメリカでも、アラビア語は学習されていますが、これをどうしたらいいのか。どういう風に勉強したらいいのか。どのくらい違うのか。というのを、日本語で試したのを教えてみたら、ざっと黒ゴチックで示したのは日本語の文語の単語ですが、38.4%でした。約4割弱でした。先ほど読んだのを聞くと全然わかんないですね。日本語には間違いないのだけれども意味がわからないですね。それが40%ある。で、アラビア語でどうなるのかとやったのが、今回の公開講座のテーマです。

その下に、エジプト映画「テロリズムとケバブ」と言う映画ですが、ちょっとどきっとしますが、これはコミックです。1993年で、コメディです。どのような話しか、その下にざっと書いておきましたが、一般市民の男性、これは写真のおじさんです。子供の健康手続きのために政府合同庁舎、これは下に写真を載せておきましたが、そこに行きます。すると職員が仕事をしないわけですね。やってくれ、やってくれと言っても仕事をしないので小競り合いとなり、駆けつけた警備員のライフルが暴発したところから、人質を取って庁舎に籠城するテロリストと誤解されてしまいます。公務員とか高級官僚とかの横暴を風刺するとともに、不満を持ちながら声を上げない国民の姿勢を風刺

したコミック映画です。10分位で編集したものをご覧いただくといいと思います。

(映画上映と部分の説明)

ちょっと止めますね。題名が「テロリズムとケバブ」なんですね。ケバブはご存知ですよ。日本で広く売っているのは、アラブ圏ではケバブと言いません。アラブ圏でケバブと言うのはひき肉を鉄串につけて焼いたものです。割と高級な料理で、下層民はちょっと奮発して買うような料理です。テロリズムとケバブが合体したのがこの映画の面白さなんです、それがここに出てきました。

(映画上映と部分の説明)

ここに題名が書いてありますが、よくわからないですよ。これがテロリズムなんです、発音がイルハーブです。ケバブはカバーブといいます。イルハーブとカバーブは韻を踏んでいますが、割りと面白い題名です。さあアラビア語はどうですか。今お聞きになったのがアラビア語ですけども。まあわかるわけは無いですよ。勉強していなければわかるはずは無いのですけれど、普通の言語です。今映画を見ましたが、もちろんしゃべっているのは口語です。ですからこの大学で1年2年勉強しても、ほぼ90%は聞き取れません。口語を勉強しない限りは無理なのです。

では問題は、口語体と文語体はどれほど違うかという問題です。先ほど日本語では40%というのを見ましたけれども、それを数えてみたのが次のページの円グラフです。図1、対応単語、どれほどのどのような単語が対応しているか。同じものが対応しているのが白い部分で、違う単語が対応しているのがねずみ色の部分です。ですから大体60から70%が同じ単語です。で、30~40%が違う単語です。右の、円グラフを見てみましょう。ほぼ同じように見えますが、こちらは異なり語数です。延べ異なり語数と頻度数と言うのはこれまでコーパスのところで解説ありましたか?延べ頻度数と書いてありますが、その中で1番良いのは、白いところ、同源語、名詞、動詞、その他、その次、alと書いてありますね。alというのはアラビア語の定冠詞です。最初の第10シーンまで数えたのですが、その中で定冠詞が104回あったのです。その次にwa、これは英語で言うとandです。「そして」とか「なにと」とかと言うandです。すると定冠詞が104回、wa, and は75回、それをそのまま数で数えるのが延べ数です。それが右側の異なり語数と言う所では1個です。定冠詞alは1個です。コーパスの中で、1万単語の中でどういう風に単語を数えるかというときに、1つの同じ単語を1度数えるのが異なり語数と言います。よく使うのが例えば104回、この104というのが延べ頻度数です。ここで調査したのがそれほど多い単語の数でない、それほど変わらないですね。ざっと6、7割が同じ単語で3、4割が異なる単語です。こういう結果が出ました。ですから先ほどの日本語の文語体を数えた結果、ちょっと文語体が多いか少ないかぐらいのほぼ同じ位です。ですからアラブ人の中でも教養のない人はニュースを聞いてもあれぐらいしかわからないのです。教養のある、高校を出たとか大学まで行ったら大体わかるといえます。そのぐらいの教養がない、小学校をロクに出ていないとか、そういう人はどのぐらいわかるかな、もちろん本が読めない、新聞も読めない、文盲もかなりいますからね。私がカイロからアレクサンドリアに行く時に、車内で切符を売る、見る人がいるのですが、その人は文盲でした。何番かと言うのをこちらが言わないといけないのですが、そんな人が案内していました。そんな人が仕事しているのですが、時々街の中でも時々会います。

その次の単語の対応関係と言う所ですが、先ほど日本語の文語は今私が話しているのと同じ発音で読みましたよね。ところがアラビア語の場合には発音が割と違うのです。割と違うといっても、慣れると同じに聞こえてくるのですが、でも最初に聞くと随分違うのです。

最初の表をちょっと見てみましょう。同源の対応、同じ単語対応していると考えてください。数字の3、タラータ、文語ではサラーサ、英語のthの発音に近いのですが、ぱっと聞いて同じ単語に聞こえますか?同じ単語だと念を押されたのでそう思い込み始めているかもしれませんが。

その次のアパート、最初のエスが上下に伸びたような音はシャという発音記号です。その次の?のようなものは声門閉鎖音ですが、皆さんわかりますか?聞いたことありますか?皆さん声門を閉鎖してみてください。できましたか?皆さんしょっちゅう使っているはずなのですが、意識的ではないので、あいうえおの「あ」を3回言ってください。声を出して言うみてくださいね。今のは声門閉鎖が少し入っています。あああ、いれないと、あああ、これは声門が閉鎖していません。違いがわかりますね? その途中の切れた感じが声門閉鎖です。声門が閉じて開いているのです。だからアパートは口語では、シャ・アです。文語では、qで書いてありますけれども、カよりもかなりかなり奥の方で発音します。カカカ。カでなくて、カ。文語では、シャ・カ。おかしな単語に聞こえますか?

「家」、これは口語ではベー、文語ではバイ。アイという母音がエーになるのはわかりますか、納得いきますか?他の語学をやっても時々出てきますよね。英語でもアイがエーになりますよね。アラビア語でももともとはバイトだったのですがこれはベーというように会話では発音します。ちなみにですね、ローマ字のアルファベットのベがありますが、あのベはこれです。

もともとのローマ字は何文字か知っていますか?ローマはギリシャから学んだのですが、ギリシャは誰から学んだか知っていますか?ご存知の方は?ご存じないですか、フェニキアです。地中海で貿易を行っていたフェニキアです。フェニキア語というのはアラビア語と姉妹関係です。アラビア語・ヘブライ語・フェニキア子は姉妹関係で、「家」という単語はほぼ同じで、ベです。あるいはバイ、ヘブライ語でもバイ、ベ。ベツレヘムというのがありますね。あれはヘブライ語ではベツ「家」、意味としては住む家ではなくて神殿ですがね。アルファベットというのはもともと家なので、家の象形文字です。それからだんだんとギリシャに行くと、Bになりました。Aはちなみにフェニキア語で「牛」です。もともと牛の頭がひっくり返ってこんな感じです。それが今こうなったのです。ですからアラビア語はそのまま覚えて使えるのです。ですからベートとかバイトとか、そういうのに慣れてくると、あ、同じ単語を使っているな一となるのですが、初めてアラビア語を勉強する人にとっては、「さっき先生バイトと言ったのにこっちではベートと言っていますよ」とか、同じ単語となかなか認識できないのですね。こういうのも同じ単語だと言うふうに考えると、6、7割は同じ単語だと言うことです。だから聞き取ろうとすると文語体と会話体は両方聞く訓練が必要になってきます。

それから下から2番目の疑問詞というのがありますね。これは同源では無いのです。エジプトではエー、文語ではマー、ぜんぜん違うでしょう。どうですかみなさん、円グラフのねずみ色のところ。このように同じような単語は6、7割、全然違う単語は3、4割ということです。全編が2時間で結構長いんですけども、さっき少し流しましたけれども、全編見たければYouTubeなどにあがっています。字幕なしですが。実は字幕つきのも私は見えています。

さあここで、こういう調査結果が出ました。皆さん、もう少しアラビア語に浸っていただくために、言葉から見える社会と言うのを少し用意しました。

まずご覧いただきたいのは最初に「田舎の...」というのがありますが、右にエジプトの地図があります。カイロの場所を確認しましょう。カイロはどこですか?ありましたか?かなりエジプトの北のほうにあります。カイロから、アレクサンドリアはわかりますか?海沿いのアレクサンドリアです。海沿いにオデッサとか、その左にアレクサンドリア。高速鉄道では無いですけど列車でノロノロとアレクサンドリアに2、3時間行く距離です。カイロから緑色に塗られている所が緑豊かな地域です。ここは農村です。逆に南に行くとナイル川の上流のほうに行くんですけど、途中エルニミヤと言う山が3つあるところがあります。そのアッシュートとソハーグ、アドビス...、それから赤い文字でルクソールがあります。このエルニミヤあたりからだんだんと上エジプトと言うところになります。

ルクソールと言うのは古代エジプトの都ですね。昔で言うテーベです。古代エジプトファンにはテーベの方が。カルナックにはカルナック神殿があって、王家の墓もルクソールにあります。エジプトの遺跡が詰まっているのはこの辺です。

緑色の塗られているところを見るとわかりますけれども、細いですよね。ナイル川は幅がわずか多くて、2、3キロの間しかない。貧しいのです。その先は砂漠でしかない。貧しくて、ださくて、田舎もので、カイロに出てきて、さっきみたいな靴磨きをして、という社会認識がある。それを如実に表しているのが、そこに書いてありますが、合同庁舎という意味なのですけども、カイロの発音では?の記号で裏返しなのですが、発音記号で喉をちょっと絞める、あーあーあー、と言うように。ムガンマー、ムガンマーという発音です。上エジプト発音だとそれがムジャンマー、カイロのムガンマーという発音をムジャンマーとすると、途端に田舎臭くなります。この田舎臭さを感じるには現地人になるしかないのですが。私も客観的に見てしまうので、ムガンマーもムジャンマーも同じ物でないかと。その靴磨きがこんなところで靴磨きをしているのかと言われて、「そうだ、この2坪を政府から買い取った」と言っている。その2坪と言うのが、?みたいのがさっきの声門閉鎖音ですね。カイロで言うと、エラーテーン、上エジプトだと、ギラーテーン、となります。カイロ発音のイと言うのをギとやりますと、途端に田舎くさいのです。日本でもそういう例があるのをお分かりですか?田舎くさい発音。例えば「エバラゲ」と言うと東北から出てきた田舎くさい出稼ぎのおっちゃんではないか、という雰囲気わかりますか?あと、シとツが区別ができないとかというのがありますね。寿司食って...寿司だかなんだかわからないと言うような、ちょっと田舎くさいと言う感じが湧いてきませんか。そういうものだと思います。こういうのは発音そのものに原因があるのではなくて、どういう人たちがそういう発音をしているのかということにあるので、発音だけ聞いてもわからない。それをちょっと聞いてみましょうね。ムガンマーと言うのとムジャンマーとギラーテーンというの聞いてください。

(動画を見る)

この映画でこの役者のこの靴磨きのこのおじさんは、本当のカイロ人です。なので、喋っているうちにカイロの発音が出てきますけど、わざとムジャンマー、と言わせたりしている。なんでそういう発音させているかというと、田舎から出てきた貧しい出稼ぎの人だという雰囲気を出そうとしているわけです。それから先ほどのあらすじには出してなかったんですが、太った自殺願望の男の人がいます。その人が遺書をお母さんに書いてもらっているのですが、それがだんだんと、文学的になってきて文語体になってくる。その中でちょっと聞き取れるかどうかかわからないですが、昇降

機、皆さん昇降機などと言わないでエレベーターと言うでしょ「でも」なんだけど、「しかし」と言うように、ワラーキンナムといたり、会話だったら、...「止まらざりし」、最後、カザルという所を、「これが運命なりや」と言ったりしてます。ミッサード、と言ったり、...その辺を味わってみたいと思います。

(動画を見る)

ちょっと聞き取るのは大変だと思いますけれども、セリフの作り方として、口語体で遺書を書くというのはわりと軽い感じですが、文語体で荘厳に遺書をかきはじめるというのが、アラビア語を知っている人には逆に滑稽に見えます。それから右の人はテロリストにされてしまった人で、職員の人ですが人質のようになっています。内務大臣との連絡をするためにトランシーバーを持ってこいと持ってこさせるのですが、誰に持ってこさせるかということ子供に持ってこさせるのです。それを受け取れと言うときに、その下の括弧3番のところですが、テロリストの威厳として、「汝の母、子を失いし」と叫べと言うのですが、それが文語体なのです。ちょっと聞いてみましょう。

(動画を見る)

いま変な声を出してますね。「ウ～」と。さっき「カバーブ」とおばさんが言ってた時も「ウ～」とやっていたのがわかりましたね。あれは女の人が喜びの時に出す声で、ザクルータといいます。エジプトだけでなくからアラブ全域で結婚式などで喜びの声で「ウ～」と出します。これもこの映画で味わえるアラブの文化だと思います。内務大臣の威厳と言うのもありますが、トランシーバーでかける時も文語体をかなり入れて文語で話すということです。

ちょっと時間もだんだん押してくるので省略しますが、もう一つこれは解説だけになりますが、エジプト方言の疑問詞位置、5番のところになりますが、エジプト・アラビア語の疑問詞「何処」とか「何」とか、文章の位置が曖昧なのです。他のアラビア語は全部文頭です。英語もそうですね。日本語は文頭に置きませんよね。後に来たり文の途中に来たり、「あなたはどこに行くの?」「どこにあなたは行くの?」、最初におかなくともいいですね。「名前は何か?」は最後に置きますが。アラビア語の中でエジプト語だけは、文頭に来ないという特徴があります。でも、それが文頭に来る時もあるので、どういう違いがあるのか、というのがあまり明らかではなかったのです。なぜ明らかではないのかと言うと、第一回の講義でもありましたが、それはコーパスを使ってないからですね。話し言葉ですからコーパスはなかなか作れないのです。書き言葉だったらいろいろな本をスキャンしてやれば良いのですが、話し言葉だったらスキャンできないでしょう。だから何万単語というのはなかなか集められないのでしょう。今までの学者だったら、「あー、こういう場合だったらどう言うかなあ」とかいろいろ自分でいくつかの例を思い浮かべて考えて行く。

今回は今のシナリオのもう一つのシナリオを使って疑問詞の位置と言うのを調べてみました。そうしたところ、動詞があって疑問詞が主語ではない場合は、元の位置にあるということがわかりました。ですから文頭には来ないのです。

基本は文頭に来ないのです。ところが次のページにあります、コンピュータ文というのがあります、ご存知ですかね。日本語で言えば「です」、英語では「Be動詞」ですが、そういうのをコンピュータといいます。コンピュータの文、なにになには何々だという文の半分は文頭にきている、ということになります。少なくとも、そうしたコンピュータ文だけが疑問詞を文頭に置いている、ということがこの

調査でわかりました。

疑似分裂文と言うのは、コピュラ文の1種です。日本語でも良いのですが、「なにになにするのは何か?」、例えば括弧10のところですが、その「君がしていることは何か?」は言い換えることができます。「君は何をしているのか?」、これは普通の疑問文です。これを作り換えて、「君がしていることは何か?」と言うのが疑似分裂文と呼んでいます。

つまり、「なにになには、なにになにか?」と言うのがコピュラ文ですね。下の表を見ると、コピュラ文、文頭に疑問詞がある文頭移動している文が、49とあります。元の位置、文頭にないものが35、動詞文は文頭移動が4つありますが、元位置が145で圧倒的に多い。一番下の疑似分裂文、これは22というのが文頭移動、と言うことで、今まで強調だから文頭に疑問詞がくるのではないか、と言われていたものは、実際の映画のシナリオを見てみると、コピュラ文だけであるということがこれで明らかになった、という報告で、この最後のエジプト方言の疑問詞というところです。

ここまでで、一通りご紹介しましたが、どうですか? コーパスというにはかなり小さいコーパスで映画のシナリオだけ、これも中心にお見せしたのが「テロリズムとケバブ」だけですが、実際にやってみたのが、もう一つ「ハッサンとマルコス」という映画をあわせて調べました。単語の文語との割合の方は、シナリオですが、それも一つのコーパスで話し言葉を文字にしてどういうものがあるか、というのを全部数えてみる、というのがコーパスです。みなさんもちょっと気になることがあったら、コーパスをやってください。ムースはどんな風になっているか、この中ではこんな発音をしているな、とか、こんな感じの使い方をしているな、とか、ドラマのシナリオをちょっとみたり小説を見たりしたり。日本語でもちょっと気になることがありますよね。「ダサイ」というのがいつ頃から出て来たのか、とか「やばい」というのがいつ頃から使い始めたのか、気になりますよね。そしたら、30年前の小説に出ているのか、そういうのを数えてみるのがコーパスの楽しさなのです。ドイツ語とか英語とか日本語の文語コーパスとか巨大なものは私など扱えないのですが、自分で手軽に試してみるというコーパスの使い方もあると思います。

以上ですが、ご質問などあればお受けします。(完)