

# Är en zigenare mer oanpassningsbar än en rom? En pilotstudie om kollokationer för orden *Cikán* och *Rom* i modern, tjeckisk tidningstext

Irene Elmerot

Institutionen för slaviska och baltiska språk, finska, nederländska och tyska

irel5167@student.su.se

*Abstract. Are “Gypsies” less adaptable than “Roma”? A pilot study on collocations for the Czech words Cikán and Rom in modern newspaper and magazine texts*

The purpose of this study is to explore the linguistic Othering of Roma people in a large corpus of modern newspaper and magazine texts in Czech. The texts that form the basis of the chosen corpus are taken from the Czech Republic’s largest daily newspapers and some of the largest magazines, which have sizable readerships. Answers to my hypotheses and questions about adjectives are given from a purely statistical perspective. Although more negative than positive adjectives are found before both studied terms – *Rom* and *Cikán*, many are neutral or stereotypically neutral. This study demonstrates how corpus linguistics can contribute to research into Othering, since the method used provides results from a substantial amount of basic data. The results are in line with previous research, and confirm it by means of an analysis of a large amount of data.

## 1. Inledning

Syftet med denna studie är att undersöka den språkliga andrafieringen av romer i ett stort textmaterial med modernt, tjeckiskt språkbruk. De tidnings- och tidskriftstexter som ligger till grund för materialet kommer från Tjeckiens största dagstidningar och några av de största tidskrifterna. De har alltså lästs av många personer när de gavs ut, men kan först numera studeras överskådligt genom att arbeta med ett korpus som innehåller dem.

Detta är en statistisk undersökning av andrafieringen av romer i det moderna språkbruket i tjeckiska. Källmaterialet är tidningstext från subkorpora till det

tjeckiska nationalkorpuset, i detta fall samlingen SYN (se 5. Källmaterial), från 1989 (enstaka inlägg) till 2009. Metoden grundas i att sälla ut adjektiv som syftar på begreppen *Rom* (svenska 'rom') och *Cikán* (svenska 'zigenare'). Resultatet av sällningen visar hur andrafieringen ser ut i de mest populära tidningarna och tidskrifterna, och denna pilotstudie kan därför ge ytterligare grund för vidare studier.

Jag vill gärna tacka lektorerna Klas Rönnbäck, för den ursprungliga idén till denna studie, och Tora Hedin för ledsagning i allt som rör det tjeckiska. Tack också till de anonyma granskarna av denna artikel för konstruktiva kommentarer.

## 1. Historisk bakgrund

Romer har funnits i det som nu är Tjeckien åtminstone från sen medeltid, ungefär lika länge som i Sverige – enligt encyklopedien *Ottův slovník naučný* sedan år 1416 (*Ottův slovník naučný* 2892:366), eller möjligen redan tidigare. Under romantiken beskrevs resande (romer eller ej) ofta som särskilt exotiska och kulturellt aktiva, och det mest kända exemplet i tjeckiskt sammanhang är Karel Hynek Máchas bok *Cikáni*, som även har satts upp som pjäs. Ett stycke ur den visar hur dessa, som ändå fått benämningen *Cikán*, 'zigenare', stereotypiskt beskrivs som underdåniga (åtminstone den yngre), resande musiker:

Byli to dva cikáni. Napřed šel malý, po maďarsku vystrojený mladík [...] Postava jeho byla nepatrná, malá; hlava, jakož i on celý, k pravé straně se nachýlela. Jeho pravá ruka byla o poznání delší nežli levá, a ze zvyku nosil obličej ustavičně do země obrácený; leč obličej jeho byl krásný, obzvláště vysoké, hrdě sklenuté čelo a temnomodré oko, jehožto smutné časem zplanutí nevolně k soucitnosti budilo. Druhý cikán, mnohem starší, byl celý modře a méně skvostně oblečen. Nad opálenými tvářemi hořely černé oči, co dvě hvězdy na večerním nebi. Mladší měl převěšený přes rameno lesknoucími se strunami potažený cimbál; starší nesl pod pažím hudební nástroj, podobný našim houslím. Byli cestující hudebníci. (Mácha 1924:14)<sup>1</sup>

Dagens beskrivningar tycks betydligt mindre romantiska, stolta och färgglada.

Romerna har en speciell ställning i Tjeckien – å ena sidan klumpas de ofta ihop med folk från andra länder (Zalabáková 2012:83<sup>2</sup>), å andra sidan nämns deras språk,

<sup>1</sup> Engelsk översättning, gjord av Geoffrey Chew, nedladdad från [http://www.academia.edu/9874209/Karel\\_Hynek\\_M%C3%A1cha\\_Cik%C3%A1ni\\_Gypsies\\_1835\\_chapter\\_2](http://www.academia.edu/9874209/Karel_Hynek_M%C3%A1cha_Cik%C3%A1ni_Gypsies_1835_chapter_2): They were two gypsies. In front came a short youth [...] His figure was slim and slight, and his head, like his whole person, was bent to the right. His right hand was noticeably larger than his left, and out of habit he kept his eyes constantly turned to the ground; but his face was handsome, an especially high, proudly arched forehead, and a dark blue eye, whose intermittent desolate intensity elicited involuntary sympathy. The second gypsy, much older, was dressed entirely in blue, and less extravagantly. His dark eyes, above sunburnt cheeks, blazed like two stars in the evening sky. The younger one had a dulcimer slung over his shoulder, with glittering strings; the older one carried under his arm a musical instrument similar to our violin. They were itinerant musicians.

<sup>2</sup> *Ti totiž byli v některých případech automaticky a takřkajíc jedním dechem zahrnuti do sdělení týkajících se „cizinců“ a jejich začasť „exotických“ kultur. Ocitli se tak po boku jiných menšin (s odlišnou kulturou), „ne-našinců“ nebo prostě „cizinců“, např. muslimů (kteří samozřejmě nemusí být vždy ne-Češi) a uprchlíků (tedy dalo by se říci cizinců par excellence) – s těmi všemi podle lidového (a mediálního) diskursu mají kromě*

romani, inte alls bland (icke-tjeckiska) språk som nämns i mediatexter tillsammans med ordet *mluvící*, enligt Fidler (2016:45–46). Där skriver hon också (Fidler 2016:56) att det måste anses märkligt att språket romani inte fanns bland de 30 främsta kollokationerna i media från 1989 till 2009, trots att det är ett av fyra officiellt skyddade språk i landet.

Ordet *Rom* började första gången användas i tjeckiskt material under andra halvan av 1960-talet, även om det blev mer vanligt efter sammetsrevolutionen, medan ordet *Cikán* har förekommit redan i de tidigast skrivna texterna som finns i det nationella korpuset. Båda orden får därmed anses vara väletablerade i tjeckisk text efter 1989, även om *Cikán* givetvis har en väsentligt längre historia och förekommer i vokabulären även hos människor födda före 1960. Värt att nämna är att inte alla i denna folkgrupp accepterar beteckningen *rom* – i Nationalencyklopedien konstateras att för vissa grupper betyder ordet istället bara ’gift man’, men att de accepterar adjektivet romsk om sitt folk. I exempelvis svenska TT-språket, en rekommendation från den språkvårdande Mediespråksgruppen till svenska journalister, står följande:

Vissa kallar sig själva romer, andra zigenare. Mediespråksgruppen har intrycket att ordet rom sprider sig i svenskan, men vi ser inte stöd för att ordet zigenare i sig skulle vara nedsättande. Att ta hänsyn till vad en intervjuperson vill kalla sig är snarare en fråga om artighet.

Motsvarande journalistisk praxis i Tjeckien kan givetvis påverka resultatet av denna studie för tabellerna med ordet *Rom*, men risken är liten. Ordföranden för den tjeckiska organisationen Romské občanské iniciativy, Emil Ščuka, ska också ha sagt i tjeckoslovakisk tv 1990: ”Cikáni je označení deklasované sociální skupiny, Romové je označení národa” (Zigenare är benämningen för en socialgrupp i utanförskap, romer är benämningen på ett folk) (Pape 2007:46).

## 2. Teori

Denna studie utgår från professor Masako Fidlers teori (Fidler 2016:38) om att titta på språket som en fingervisning om möjliga sinnebilder av ”De andra” som – mer eller mindre undermedvetet – ligger till grund för hur majoritetsbefolkningen ser på en viss minoritet i landet. Språkliga mönster är mer eller mindre normala för ett språks läsare, och det går med statistiska metoder att få fram vilka mönster som är vanligast, vilket också innebär att vi kan se vad en läsare bör tycka är självklart och inte lika självklart. Hon hävdar också (2016:52) att ordfrekvensen i ett välbalanserat referenskorpus (såsom SYN här) bör återspegla allmänna språkanvändningsmönster, vilka hänger samman med vad de som kan korpusets språk normalt läser i vardagen.

---

„jinakosti“ společnou „tradiční“ kulturu (o problematice romské tradice viz Hirt a Jakoubek 2004, Svoboda 2006). (Zalabáková 2012, 83)

Som Fidler också skriver i en artikel med Václav Cvrček, påverkas vi och hela vår världsbild av ordkombinationer som ofta förekommer tillsammans:

We anticipate that such expectations are inevitably built on patterns of language use through consistent exposure. (This idea is not entirely new; other scholars have made observations that point in this direction.) Take, for example, the phrase “illegal immigrant”. The two words prime one another through repetition; readers exposed to this phrase are likely to think of this sequence even when encountering the word immigrant without a modifier; for such readers the word often invokes suspicion about migration (Stubbs 1996:197). Language use thus leads to the reader’s conceptualization of the world (point of view). (Fidler och Cvrček 2015:198)

### 3.1 Frågeställningar

Följande frågor kommer att ställas på materialet:

- (1) Vilka adjektiv är vanliga som bestämmningar till orden *Cikán* och *Rom* i detta källmaterial, och i vilken utsträckning är dessa adjektiv positivt eller negativt laddade?
- (2) Skiljer sig språkbruket åt mellan texter som använder ordet *Cikán* respektive *Rom*?
- (3) Finns det en mångfald av olika adjektiviska begrepp, eller är det bara ett fåtal begrepp som återanvänds?

Dessa frågor ska besvaras genom den statistiska metoden här, och de kan återanvändas även i större studier framöver. Mot bakgrund av tidigare forskning (se nedan) bör svaret på fråga 1 luta åt att negativt laddade adjektiv är de vanligaste, eller åtminstone de vanligaste utöver neutrala adjektiv. Fråga 2 torde därför besvaras med ett ”ja”, frågan är i vilken utsträckning, och huruvida det blir positiva, neutrala eller stereotypa neutrala adjektiv som får högst antal förekomster. Detta leder till följande hypoteser:

### 3.2 Hypoteser

- (1) a) att neutrala adjektiv som rör individer (*stor, kort, lokalboende* etc.) är vanligt förekommande, och  
b) att bland icke-neutrala adjektiv är negativt laddade adjektiv de mest frekvent förekommande.
- (2) att skribenter som använder begreppet *Rom* även visar mer respekt i övrigt och därmed är mindre öppna för negativa och stereotypa i sina formuleringar.
- (3) att ett fåtal begrepp återanvänds.

### 3. Tidigare forskning

Andrafiering förekommer oftare i statsvetenskapliga, medievetenskapliga och filosofiska artiklar än i språkvetenskapliga, men det finns några som på senare år har forskat just på hur romer behandlas i skrift i länderna som en gång låg bakom järnridån. Det har även blivit ett populärvetenskapligt debattämne, i skrivande stund i Sverige senast genom journalisten Hynek Pallas bok *Oanpassbara medborgare*, där

han har intervjuat människor i Tjeckien och studerat arkivmaterial om vad som har skrivits om romerna i Tjeckien tidigare. Begreppet *språklig andraftering* verkar inte i dagsläget förekomma på svenska, men däremot finns några som tidigare har använt det på engelska, bland andra Harrison (2009) samt Dako och Yitah (2012).

McGarry (2014) har undersökt romers (re)presentation i medier runtom i Europa, och konstaterar att romernas politiska identitet beror på både hur majoritetsbefolkningen upplever dem och på forskning. Han har bland annat observerat beskrivningar som ”socialt eftersatta”, och anser (McGarry 2014:769) att den huvudsakliga utmaningen är att motbevisa svartmålningen av romer som grupp i lokala, rikstäckande och internationella sammanhang.

Från Institutet för mediastudier kom en bok om den svenska mediebildens av romer (Pollack et al. 2015) där Ester Pollack och Anna Roosvall skriver om rapporteringen i Dagens Nyheter och Aftonbladet 1995–2014, det vill säga ungefär samma tid som denna studie berör. De har studerat rubrikerna till 502 artiklar, de flesta publicerade efter 2010 – detta på grund av den då ökade inflyttningen av romer från bland annat Bulgarien och Rumänien. De konstaterar (Pollack et al. 2015:21) att användningen av begreppet *rom* ökar kraftigt under perioden, men att begreppet *zigenare* ändå förekommer ända in på 2010-talet, om än oftast när artikeln handlar om historiska händelser som andra världskriget. I samma bok skriver Elitsa Ivanova om beskrivningen av romer i bulgariska medier och konstaterar att forskningen om hur romer representeras i olika länders medier inte är så omfattande (Pollack et al. 2015:73), men att de beskrivningar som har studerats är ”klichéartade och de är antingen exotiserande” – det som i föreliggande artikel kallas för stereotyp neutrala.

I en artikel jämför Catalano (2013) beskrivningen av romer i brottsrelaterade tidningsartiklar med beskrivningen av företagsledare inom finansvärlden i motsvarande, brottsrelaterade tidningsartiklar i USA och Kanada. Hon använder diskursanalys och kognitiv lingvistik för att konstatera att medvetenheten måste höjas om hur dessa media beskriver människor med olika bakgrund, men grundar sin artikel på bara 35 nyhetsartiklar. Hon kommer dock fram till liknande resultat som föreliggande studie, nämligen att kollokationerna med orden *gypsy* och *roma* bildar stereotyper och negativa tankegångar.

Tack vare att det tjeckiska nationalkorpuset är så stort, har några språkforskare börjat använda det i liknande syften, inte minst Masako Fidler. Här ligger dock fokus på de tjeckiska orden, adjektiven, till skillnad från Fidlers undersökning om hur andra språk än tjeckiska uppfattas.

Tess Slavičková och Peter Zvagulis skriver i sin artikel (2014) om hur de har arbetat långsiktigt med att beskriva så kallat hate speech i såväl brittiska som tjeckiska nyhetsartiklar, och att det mesta sådant de hittade i tjeckisk text berörde romer och vietnameser. De noterar (2014:159) att begreppet *nepřizpůsobivý občan*, ’oanpassningsbar medborgare’, är den senare tidens kodord för romer. De analyserar

här främst en enda nyhetsartikel, men skriver (2014:167) att de framöver ska arbeta med det tjeckiska nationalkorpuset. I skrivande stund hittar jag dock inte någon nyare artikel av dem som grundar sig på det.

I inledningen till boken *Obráz Romů v středoevropských masmédiích po roce 1989* (Karhanová, Homoláč och Nekvapil 2003:3) säger författarna att de med boken ville sammanställa vilka (retoriska) medel, metoder och genrer som typiskt används när journalister skriver om romer. De artikelförfattare som i boken berör Tjeckien utgick dock ifrån talat språk och tidningsartiklar om muntliga tv-debatter, samt intervjuer med tjecker om romer apropå händelserna i nämnda tidningsartiklar. Deras tjeckiska material handlar enbart om tre våldsdåd som begicks i början av 1990-talet, som alla ledde till att en rom dog. Deras arbete har gjorts med en helt annan metod än föreliggande studie, och deras källmaterial är mycket mindre. I artikeln ”Je trestné topit Roma? (Prezentace násilné smrti Tibora Danihela v českém tisku)” (”Är det olagligt att mörda en rom?” En presentation av Tibor Danihels våldsamma död i den tjeckiska pressen) skriver Anna Šabatová, Jiří Homoláč och Kamila Karhanová att där en rom blir våldsoffer, får journalister en ursäkt att skriva om negativa stereotyper, vilket avspeglas även i relativt neutrala texter. Att forskarna här fokuserar på våld innebär att romerna automatiskt får en offerroll, vilket vinklar resultatet på ett sätt som skiljer sig från mitt:

Událost, kdy se Rom stal obětí trestného činu, se tak paradoxně (při aktivaci relačního páru Romové – majorita „logicky“) stává příležitostí k tematizaci problému romské kriminality a k mediální produkci a reprodukci negativních stereotypů o Romech. Zkušenost majoritní společnosti s kriminálním a obecně asociálním chováním Romů slouží k vysvětlení (omluvě) jejich protiromských postojů i v textech, které jsou vůči romské menšině v zásadě vstřícné. (Šabatová, Homoláč och Karhanová 2003:101)<sup>3</sup>

Deras resultat visar dock på en viktig aspekt av andrafieringen, vilket gav mig ännu en anledning att undersöka det i denna studie.

Även Kluknavská och Zagibová (2013) har studerat våldsdåd, i deras fall 2011, och nämner också adjektivet *nepřizpůsobivý* (’oanpassningsbar’), samt de negativa adjektiven *rozzuřený* (ilsken) och *agresivní* (’aggressiv’). Deras källmaterial är artiklar ut tidningarna Blesk, Mladá Fronta Dnes och Právo från den 8 augusti till 8 september 2011 (2013:306).

Inför denna studie studerades bland annat också antropologen Zalabáková (2012), som bland annat skriver att så mycket av de exotiska beskrivningarna i tjeckiska media utgörs av kulturuttryck, framför allt mat och musik (inklusive dans).

<sup>3</sup> Översatt på följande vis i den engelska sammanfattningen till artikeln: *A case in which a Romany is a victim of a crime therefore paradoxically (although in terms of the Romanies – majority pairing this may be logically) offers an opportunity for discussing the problem of the Romany crime rate and for the media to produce and reproduce negative stereotypes of Romanies. The majority population’s experience with the criminal and generally anti-social behaviour of Romanies serves to explain (excuse) their anti-Romany attitudes, even in texts that are basically favourable to the Romany.*

Det finns fler som skrivit om andrafiering av romer, till exempel Doubek, Levínská och Bittnerová (2015), men deras forskning visade sig inte tillföra mer än allmän bakgrundsinformation för denna studie.

## 4. Källmaterial

SYN är ett icke-referenskorpus med fem miljoner lemmata som består av texter från alla synkroniskt skrivna referenskorpusar i SYN-serien, det vill säga sex olika korpusar. De (i skrivande stund) senaste texterna lades till i januari 2014. De flesta texter kommer från tidningar och tidskrifter, tack vare att cirka 1,1 miljarder ord uteslutande är tidnings- och tidskriftstext från 1989 till 2009. Tre av de korpusar som ingår i SYN är balanserade med andra sorters texter, de andra är enbart från tidningar och tidskrifter. Bland annat ingår alla nummer av de största dagstidningarna i landet (Mladá fronta DNES, Lidové noviny, Právo, Hospodářské noviny och Blesk) samt tidskrifterna Reflex, Respekt och Týden från 1998 till 2009. Läs gärna mer på <http://wiki.korpus.cz/doku.php/en:cnk:syn> (på engelska).

Källmaterialet omfattar stora mängder ord och kan sökas i relativt enkelt. Det finns även en så kallad wiki som dels förklarar hur sökningar kan göras, dels innehåller en kurs i hur användarna kan göra sina sökningar.

## 5. Metod

Statistiska metoder är användbara bland annat för att snabbt gå till ett stort urval som underlag. Att statistiskt titta på framför allt positiva och negativa adjektiv i denna position ger ett underlag till hur andrafieringen ser ut i den offentliga, skrivna diskursen. Här kan vi se kontrasten mellan värdeladdade ord och neutrala ord i ett stort urval text ur de vanligaste tidningarna och tidskrifterna i Tjeckien.

Fidler och Cvrček har skrivit att genom statistiskt arbete minskar risken för att forskaren är eller anses subjektiv, och statistiken kan användas för att komplettera mer kvalitativa studier (Fidler och Cvrček 2015:199). Å andra sidan nämner till exempel Watson (kommande 2017:8) att det förekommer kritik mot att använda kvantitativ metod för att undersöka uttryck, eftersom den exakta ordalydelsen kan variera – något som upptäcktes även i föreliggande studie och beskrivs nedan. Resultatet blir ändå ett smakprov på vad statistiskt arbete kan åstadkomma och visa på inför eventuella framtida, större studier.

Denna studie kommer att avgränsa sig till perioden 1989 till 2009. I sin avhandling från Oakland university skrev A.A.G. Schneeweis att det i Europa först under 1990-talet för första gången började pratas och skrivas mer seriöst och respektfullt om romer, inte bara om ”zigenare”:

Europe in 1990 understood the implications of Holocaust-like and ethnic cleansing projects; yet it also understood Gypsy otherness. In order to allow the Gypsy to be Roma, the press (among other key institutions) needed to start building a vocabulary and framework for living with difference (an ongoing process to this day). The discourse of the civilized Gypsy began carving this body of knowledge. (Schneeweis 2009:147)

Det stärker valet av period som källmaterial till en sådan här pilotstudie. Det är möjligt att användningen av orden *Rom* och *Cikán* har ändrats på fler sätt över den tid som studien täcker. Detta kan i sig vara ett ämne för framtida forskning.

I denna artikel undersöks andrafiering från en språkvetenskaplig vinkel, genom en statistisk undersökning av kollokationer, i detta fall de adjektiv som var kollokat till de två valda begreppen för samma folkgrupp. Genom sökning i korpusar (serien SYN på [www.korpus.cz](http://www.korpus.cz)) framgår vilka adjektiv som syftar på ordet *Cikán*, dels ordet *Rom*. I praktiken avgränsas sökningen till att enbart fokusera på de adjektiv som står närmast till vänster om orden, för att få fram ett meningsfullt statistiskt resultat där adjektivet med mycket hög säkerhet kan antas syfta på dessa substantiv, och inte på något annat ord i texten.

Med denna metod sorteras alltså ett antal träffar bort som är nog så relevanta rent sociolingvistiskt för den som vill se en texts sammantagna tendens och vad den egentligen säger om romer i sin helhet. Adjektiven som står i samma mening, men längre bort från substantivet i fråga, kan mycket väl också syfta på något som tillhör romerna eller på något som nämns i samband med dem, men då faller det alltså bort i gallringen (ett exempel: ”V České republice jsou skinheadi, kteří říkají o Cikánech *špatné věci*”, ur *Mladá Fronta Dnes* den 28 augusti 1997). Likaså bortfaller även exempel där det är substantiv som definierar romerna, som i ”... zbavit se přezíravosti a rozšířeného předsudku že každý *Cikán* je *nevdělanec, lenoch a kriminálník*...” (*Mladá Fronta Dnes* den 13 augusti 1997) eller ”*Cikáni jsou podvodníci*” (*Mladá fronta Dnes* den 5 augusti 2000). Det förefaller osannolikt att detta bortfall skulle göra någon större skillnad för resultatet då det inte torde finnas någon anledning att anta att dessa bestämmningar systematiskt skiljer sig åt från de adjektiv som står i direkt anslutning till substantiven.

Ett problem med metoden är dock att det bland de positiva adjektiven före ordet *Cikán* förekommer ord i negativ kontext. Ett exempel är *dobrý*, ’bra, god’ i den synnerligen negativa kontexten ”En död zigenare är en bra zigenare”. Likaså förekommer adjektivet *chytrý* (’smart’) i en sats som följs direkt av ”för sådana [dvs. smarta zigenare] måste ju finnas” – en icke-positiv mening i sammanhanget. I sin helhet är dock detta inte av avgörande betydelse för resultaten.

Adjektiv som isolerat kan förefalla neutrala eller positiva, men i praktiken är mer eller mindre stereotypiserande, förekommer också. Ett exempel är *kočovní* (’nomadisk, vandrande’). Ibland kan också texternas författare använda ett positivt begrepp för att i praktiken – mer eller mindre underförstått – antyda eller anspela på en mycket negativ stereotyp, som när det uttryckligen står att en rom är anständig,



eftersom stereotypen är att hen inte är det. Detta innebär ett visst metodproblem, eftersom sådana adjektiv blir svårare att kategorisera och sortera rent statistiskt.

Den statistiska bearbetningen av materialet bygger på att sortera ut de adjektiv som är värdeladdade respektive neutrala, dvs. positivt laddade adjektiv (kategoriserade med ett P i det ur databasen utlyfta materialet, hädanefter kallade positiva adjektiv), negativt laddade adjektiv (Neg, hädanefter kallade negativa adjektiv) och stereotypiserande neutrala (SN) respektive till synes neutrala (Neu) sådana. Under sorteringen framstod geografiska benämningar (som *tjeckisk* eller *Kolín-boende*) väldigt tydligt, så kategorin geografiskt adjektiv (G) fick läggas till. Eftersom jag inte stötte på någon tidigare modell för sådan kategorisering under arbetet, är indelningen helt och hållet min egen. Kategoriseringen återfinns i sin helhet i appendix till denna artikel.

Det första sökordet blev *Cikán* (som här kommer att översättas med det svenska ordet zigenare) i olika former, med närmast stående adjektiv markerade, och det andra sökordet blev *Rom*. Sökningen gjordes i den del av korpusen som gav träffar från åren 1989–2009 genom en lemmasökning i publicistiskt material (alltså ingen skönlitteratur). Sökningen utgick från meningar där det finns ett adjektiv inom ett ord från huvudordet. Resultatet blev 707 träffar för *Cikán* och 142 för *Rom*.

Träffar exkluderades från urvalet efter manuell genomläsning, ifall texten handlade om filmen/musikalen *Cikáni jdou do nebe*, om pjäser och filmer regisserade av Miroslav *Cikán*, om Máchas bok/pjäsa *Cikáni* eller Pusjkins bok som på tjeckiska också heter så, samt i viss mån de som handlade om fotografen Josef Koudelkas foton av romer. Därtill kunde träffarna som syftade på klubben med namnet U *Cikána* i Hradec Králove och på övriga personer med efternamnet *Cikán* sorteras bort. Kvar blev 484 träffar för *Cikán*.

Väl inne i databasen sorterades de på frekvens och de lemmata som står närmast till vänster om huvudordet valdes först ut. Ett problem med korpusens träffrapportering av lemmata blir att negerade ord inte noteras separat, utan för att kontrollera om ett ord står i sin positiva eller negativa form måste forskaren återgå till listan och se efter där. Så är till exempel fallet med ordet *přizpůsobivý*, oanpassningsbar, som har sex träffar i frekvenslistan, men samtliga dessa är egentligen negerade, det vill säga oanpassningsbar, eller oanpassbar, som Pallas (2016) kallar det. Värt att notera är också att korpussökningen inte ger träffar på feminint böjda varianter, det vill säga att inga former av ordet *Cikánka* är med bland träffarna.

## 6. Resultat

För båda orden utgör de geografiska adjektiven ungefär en tredjedel (29,7 % för *Cikán*, 36,7 % för *Rom*). De negativa adjektiven är dubbelt så många för ordet *Cikán* som för ordet *Rom* (15,8 % för *Cikán*, 8,25 % för *Rom*). De neutrala adjektiven är å

andra sidan nästan dubbelt så många för ordet *Rom* som för ordet *Cikán* (25,15 % för *Cikán*, 47,7 % för *Rom*). Värt att notera är att de positiva adjektiven är flera gånger så många för ordet *Cikán* (6,5 % för *Cikán*, 0,9 % för *Rom*), se vidare diskussion nedan. De adjektiv som hamnade i kategorin Stereotyp neutrala blev en dryg femtedel av alla adjektiven före ordet för zigenare, medan det bara blev sex procent före ordet rom (22 % för *Cikán*, 6,4 % för *Rom*).

Varken i tabell 1 eller 2 är alla ord med, utan bara de vanligaste, för att ge en mer översiktlig bild. Se appendix för en mer komplett lista.

Geografiska	Negativa	Neutrala	Positiva	Stereotyp neutrala
rumunský 23	špinavý 14	mladý 14	slavný 9	kočovný 25
český 14	přizpůsobivý 6	olašský 9	dobrý 8	potulný 15
mořský 13	mrtvý 4	představený 8		kočující 12
místní 12	zatracený 3	starý 6		slušný 11
španělský 10		mnohý 5		černý 6
maďarský 7		jediný 4		bílý 5
zdejší 6		nazvaný 4		
andaluský 6		obyčejný 4		
moravský 6		další 4		

Tabell 1. Före *Cikán*. Siffran motsvarar antalet förekomster av sammanlagt 484 träffar.

Geografiska	Negativa	Neutrala	Positiva	Stereotyp neutrala
slovenský 8	opilý 1	mladý 6	zlatý 1	slabý 2
český 7	tokající 1	místní 5		zaměstnaný 2
kostelecký 3	brutální 1	olašský 4		kočující 1
vsetínský 3	přizpůsobivý 1	jiný 3		sebejistý 1
moravský 3	nepřátelský 1	starý 2		žebrající 1
rumunský 2	rozvášněný 1	jediný 2		
francouzský 2	vyhlížeující 1	ostatní 2		

Tabell 2. Före *Rom*. Siffran motsvarar antalet förekomster av sammanlagt 109 träffar.

	Geografiskt adjektiv	Negativt adjektiv	Neutralt adjektiv	Positivt adjektiv	Stereotypt neutralt adjektiv
Före <i>Cikán</i>	29,75 %	15,9 %	25 %	6,6 %	22,5 %
Före <i>Rom</i>	36,7 %	8,25 %	47,7 %	0,9 %	6 %

Tabell 3. Procent av respektive tabells träffar.

Det enda ord bland dessa som förekommer i negerad form är adjektivet *přizpůsobivý*, det vill säga 'anpassningsbar', som dessutom enbart förekommer negerat – vilket alltså bevisar påståendena att oanpassningsbar är ett vanligt förekommande adjektiv i samband med det tjeckiska ord som motsvarar det svenska *zigenare*, och att det förekommer en gång ihop med ordet *Rom*.

Frågeställningarna besvaras på följande vis:

Som tabellerna visar, är de allra vanligaste adjektiven rent geografiska och neutrala ord. Dock är de rent negativa fler än de positiva både framför *Cikán* och *Rom*, vilket framgår av tabell 3. Värt att notera är att andelen positiva ord förefaller vara fler runt användandet av ordet *Cikán*, men *dobry* ('bra') används i hälften av fallen inte i en positiv kontext.

Det är en mångfald ord som används, eftersom inget kommer upp i mer än 5,6 procent av alla ord som står närmast varken *Cikán* eller *Rom*. Det vanligaste negativa adjektivet är dock *špinavý*, 'smutsig', som förekommer i 1,9 procent av fallen, och bara i samband med ordet *Cikán*. De olika adjektiv som betyder något i stil med icke-bofast (kategoriserat som stereotypt neutralt) förekommer i 7,3 procent av fallen före ordet *Cikán* och bara i ett enda fall (0,7 procent) ihop med ordet *Rom*.

Hypoteserna besvaras på följande vis:

**1. a) att neutrala adjektiv som rör individer (*stor, kort, lokalboende etc.*) är vanligt förekommande.**

Ja, de neutrala adjektiven är väldigt vanligt förekommande, och vill en slå ihop dem med de geografiska och/eller med de stereotypt neutrala, blir det än mer så. Det är tydligast för tabell 1. Det är något förvånande, på ett för mänskligheten positivt sätt, och är en motvikt till den forskning som framhäver att det så ofta förekommer negativa aspekter i artiklar om romer.

**b) att bland icke-neutrala adjektiv är negativt laddade adjektiv de mest frekvent förekommande.**

För tabell 1 (före *Cikán*) stämmer denna hypotes bra. Det vanligaste av de negativa adjektiven, *špinavý*, är lika vanligt som ordet *mlady* ('ung') bland de neutrala eller *český* ('tjeckisk') bland de geografiska adjektiven. För tabell 2 (före *Rom*) förefaller det också stämma, men antalet värdeladdade adjektiv är markant lägre. Före ordet

*Cikán* förekommer positiva ord inte ens hälften så ofta som negativa (32 förekomster, inklusive fyra där *dobrý* egentligen skrivits i en negativ omgivning). Före ordet *Rom* finns bara ett enda positivt ord, jämfört med nio negativa.

## **2. att skribenter som använder begreppet *Rom* även visar mer respekt i övrigt och därmed är mindre öppett negativa och stereotypa i sina formuleringar.**

Ja, detta verkar stämma. En tiondel av förekomsterna (8 + 6) rör någon som är ung eller slovak – adjektiv som inte kan anses nedsättande i sig, men som självklart kan användas i nedsättande sammanhang. Bland de negativa adjektiven förekommer inte *špinavý* över huvud taget, men det finns en förekomst av *nepřizpůsobivý*, 'oanpassningsbar'. Värt att notera ur ett samhällsperspektiv är dock att mindre öppen rasism inte nödvändigtvis betyder mindre rasism – kanske har det bara blivit lite mindre accepterat, i linje med det Schneeweis skriver.

## **3. att ett fåtal begrepp återanvänds.**

Detta stämmer delvis – ordet för smutsig förekommer 14 av 77 gånger (dvs. i 18 procent av fallen) före ordet *Cikán*, och ord som betyder något i stil med vandrare eller kringresande förekommer 52 gånger av 109 (i drygt 47 procent av fallen). I övrigt förekom många olika ord, och *oanpassningsbar* förekom inte så ofta som forskningen och populärvetenskapen kan ge intryck av.

## **7. Slutdiskussion**

Kvantitativa studier på korpusar kan fortfarande kännas nytt inom språkvetenskapen, även om det var flera årtionden sedan forskare slutade påstå att det inte skulle gå att utforska det Saussure kallade för *la parole*, användningen av språket i konkreta situationer. Mitt arbete har utgått från miljontals ord, ur vilka har plockats de som faktiskt står närmast varandra, som förekommer tillsammans och därmed kan pekas ut som vanliga kollokationer. Metoden har vissa brister, men låter oss ändå tydligt se vad som är vanligast och vilka mönster som förekommer i ett stort korpus. Resultatet stämmer med mycket tidigare forskning, men visar på andra perspektiv och andra sätt att mäta det språk som folk får sig till livs i vardagen, när de läser de vanligaste tidningarna och tidskrifterna.

Hypotes 1 a) stämmer delvis. Bara 9 (före *Rom*) respektive 22 procent (före *Cikán*) av adjektiven är uttalat värdeladdade, och en förvånansvärt hög andel är till synes neutrala. Med den diskriminering som vi vet att romer utsätts för var förväntningen snarast att fler skulle vara värdeladdade. Metoden gör att vi dock inte kan utläsa hur artikeln som adjektiven kommer från är skriven, och resultatet skulle kunna kompletteras med en större artikelstudie om hur texterna som helhet är formulerade.

Hypotes 1 b) stämmer enligt resultatet. Detta är i linje med vad andra forskare har konstaterat – att romer beskrivs negativt och med en tydlig andrafiering i tjeckisk media. Det är tydligare att kontextlöst positiva ord, som *dobry*, står i en negativ kontext än att kontextlöst negativa ord i själva verket står i en positiv kontext.

Hypotes 2 stämmer, vilket bland annat procentsatserna för negativa respektive neutrala adjektiv före orden *Cikán* respektive *Rom* visar (negativa före *Cikán* = 16 procent, före *Rom* 8 procent). En romsk tradition som ofta beskrivs av forskare som själva är eller har umgåtts mycket med romer (se Vivian och Dundes 2004 samt Tesär 2012), är de renlighetsideal som finns, och som innebär strikta regler för hur och att människor, kläder och husgeråd ska tvättas. En paradox i resultatet här är då att det negativa ord som är mest högfrekvent ihop med ordet för zigenare är *špinavý*, 'smutsig'. (Adjektivet förekommer inte heller alls bland dem som använder ordet rom.) Detta är en stereotyp som återkommer i alla möjliga sammanhang, trots att det samtidigt motstrider romernas renhetsideal. Schneewis tar i sin avhandling (2009:75) upp flera exempel på hur vanligt detta har varit genom tiderna, från 900-talets Persien till tyska krönikor, ryska ordspråk och brittisk tidningstext från 1960-talet, och det förekommer än idag.

Hypotes 3 stämmer delvis, på så sätt att orden för smutsig och kringresande är vanliga, men i övrigt är det ganska stor spridning, framför allt före begreppet *Rom*.

Inledningsvis fanns ingen förväntan på att så många adjektiv skulle vara geografiska. De fick till slut bli en egen kategori, eftersom de inte per definition måste vara neutrala, även om de ser sådana ut. Att just *tjeckisk* är det näst vanligaste (före ordet *Cikán* är *rumänsk* vanligare, före ordet *Rom* är *slovakisk* vanligare) av de geografiska adjektiven leder till frågan om tjeckiska romer på något sätt framställs som mer eller mindre positiva än andra. Den får besvaras i en annan studie. Möjligen kan en anledning vara att många, även yngre, romer fortfarande anses vara från Slovakien eller andra länder, eftersom många romer kom därifrån under arbetskraftsinvandringen efter 1945. Det tredje vanligaste före *Cikán* är dock *mořský* ('havs-'), vilket kommer från artiklar om andra länder. En fråga som uppstår är om detta fenomen, att de geografiska adjektiven är så många, är ett tecken på "vi och dem"-mentalitet? Nu är romer en folkgrupp som förekommer över stora delar av världen, så det kan vara en tillfällighet, men det är svårt att bevisa.

Till skillnad från Zalabákovás resultat återfanns i denna studie inget typiskt kulturellt uttryck, eftersom det bara finns någon enstaka förekomst av adjektiv som har med musik att göra, och ingen som rör mat (just mat nämner inte heller Zalabáková i samband med romer, det förekommer oftare tillsammans med andra grupper). Hur användningen av de båda begreppen har utvecklats över tid är en annan intressant fråga, men inget som passar denna studie.

Framöver skulle resultatet av denna studie kunna utgöra grunden för ett djupare arbete där dessa adjektiv och uttryck kan jämföras med adjektiv som används även om andra icke-tjeckiska minoritetsgrupper i republiken Tjeckien, för att kunna undersöka om det är en allmän andrafiering som pågår, som inte är specifik för just romer. Syftet med ett sådant arbete skulle vara att motsäga dem som hävdar att romer är på ett visst sätt genom att peka på att andra grupper omtalas eller har omtalats på samma sätt. Att jämföra hur dessa begrepp har använts diakront, vilka uttryck som har varit mest använda under olika tidsepoker, verkar inte ha gjorts förut för detta källmaterial.

## Litteratur

- Catalano, Theresa. 2013. "The Roma and Wall Street/CEOs: Linguistic construction of identity in U.S. and Canadian crime reports." *Faculty Publications: Department of Teaching, Learning and Teacher Education. Paper 182*.  
<http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1181&context=teachlearnfacpub>, 2016-09-14.
- Dako, Kari, & Yitah, Helen. "Pidgin, 'broken' English and Othering in Ghanaian literature." *Legon Journal of the Humanities* 1; 202–230.
- Doubek, David, Levínská, Marketa & Bittnerová, Dana. 2015, "Roma as the Others", *Intercultural Education*, vol. 26, no. 2, s. 131.
- Fidler, Masako & Cvrček, Václav. 2015. "A data-driven analysis of reader viewpoints: Reconstructing the historical reader using keyword analysis." *Journal of Slavic Linguistics* 23(2), 197–239.
- Fidler, Masako. 2016. "The others in the Czech Republic: Their image and their languages." *Multilingualism and Minorities in the Czech Sociolinguistic Space, a special issue of the International Journal of the Sociology of Language* 238, 3; 37–58.
- Harrison, Gai. 2009. "Language Politics, Linguistic Capital and Bilingual Practitioners in Social Work". *The British Journal of Social Work*, 1/1/2009, Vol. 39, 6; 1082–1100.
- Karhanová, Kamila, Homoláč, Jiří & Nekvapil, Jiří. 2003. *Obraz Romů v středoevropských masmédiích po roce 1989*. Brno: Doplněk.  
[http://ulug.ff.cuni.cz/projekty/romove/obraz\\_romu.pdf](http://ulug.ff.cuni.cz/projekty/romove/obraz_romu.pdf), 2016-09-14.
- Kluknavská, A, & Zagibová, L 2013. "Neprispůsobiví Rómovia a slušná väčšina? Spravodajský diskurz po násilných udalostiach na severe Českej republiky 2011". *Central European Political Studies Review / Stredoevropske Politicke Studie*, vol. 15, no. 4, s. 300–323.
- Křen, M., Čermák, F., Hlaváčová, J., Hnátková, M., Jelínek, T., Koček, J., Kopřivová, M., Novotná, R., Petkevič, V., Procházka, P., Schmiedtová, V., Skoumalová, H., Šulc, M. 2014. *Korpus SYN, version 3 from 27 Jan 2014*. Prag: Ústav Českého národního korpusu FF UK. <http://www.korpus.cz>.
- Mácha, Karel Hynek. 1924. *Cikáni a jiná próza*. Prag: Nakladatel B. Kocí, Knihkupec Praze I., Masarykovo Nábřeží č. 14.  
<http://kramerius4.nkp.cz/search/i.jsp?pid=uuid:287d5e20-e7d2-11e3-b72e->

- 005056827e52#monograph-page\_uid:23be7510-22ef-11e4-8c14-5ef3fc9bb22f, 2016-09-14.
- McGarry, Aidan. "Roma as a political identity: Exploring representations of Roma in Europe". *Ethnicities*, vol. 14, no. 6, ss. 756–774.
- Nationalencyklopedin, "romer".  
<http://www.ne.se/uppslagsverk/encyklopedi/lång/romer> (hämtad 2016-11-07)
- Ottův slovník naučný: illustrovaná encyklopaedie obecných vědomostí. Pátý díl.* (1892). Praha: J. Otto.  
<https://ia802608.us.archive.org/35/items/ottvslovnknauni55ottogoog/ottvslovnknauni55ottogoog.pdf>, 2016-09-14.
- Pallas, Hynek. 2016. *Oanpassbara medborgare: historien om förföljelsen av de tjeckiska romerna*. Stockholm: Atlas.
- Pape, Iveta. 2007. *Jak pracovat s romskými žáky*. Prag: Slovo 21
- Pollack, Ester, Roosvall, Anna, Caldaras, Hans, Szoppe, Adam, Hammarberg, Thomas, Ivanova, Elitsa & Lars Truedson. 2015. *Mediebild av romer*. Stockholm: Institutet för mediestudier.
- Schneeweis, Adina Alexandra Giurgiu. 2009. *Talking difference: discourses about the Gypsy / Roma in Europe since 1989*. (Diss.) <http://hdl.handle.net/11299/56088>, 2016-09-14.
- Slavíčková, Tess, & Zvagulis, Peter. 2014, "Monitoring anti-minority rhetoric in the Czech print media: A critical discourse analysis". *Journal of Language & Politics*, vol. 13, no. 1, ss. 152–170.
- Tesár, C. 2012. "Becoming Rom (male), becoming Romni (female) among Romanian Cortorari Roma: On body and gender." *Romani Studies* 22.2, 113–140, <https://muse.jhu.edu/article/489108/pdf>, 2016-09-14.
- TT-språket. 2013. "Omdiskuterade ord", <https://info.tt.se/tt-spraket/navigation/rekommendationer-fr-mediespraksgruppen/omdiskuterade-ord/>, 2016-11-04.
- Vivian, C. and L. Dundes. 2004. "The Crossroads of Culture and Health Among the Roma (Gypsies)." *Journal Of Nursing Scholarship*, 36.1, 86–91, <http://www.ncbi.nlm.nih.gov/pubmed/15098424>, 2016-09-14.
- Watson, Christine. Kommande 2017. "Truth, doubt and hearsay in 17th-century Russian news translations". Utkast skickat av författaren, läst 2016-09-15.
- Zalabáková, Jitka. 2012. Afričané, Romové a další „cizinci“ v české mediální realitě. *AntropoWebZin* 2/2012, 79–87.  
<http://www.antropoweb.cz/media/document/antropowebzin-2-2012-final.pdf>, 2016-09-14.